Psychological Bulletin

CONTENTS

Published Bimonthly by the
American Psychological Association

55, No. 5 September, 1958

The Psychological Bulletin contains evaluative reviews of research literature and articles on research methodology in psychology. This Journal does not publish reports of original research or original theoretical articles.

Manuscripts should be sent to the Editor [illegible] Harry Helson, Department of Psychology, University of Texas, Austin, [illegible]

Preparation of articles for publication. Authors are strongly advised to follow the general directions given in the "Publication Manual of the American Psychological Association" (19[illegible] Revision). Special attention should be given to the section on the preparation of the references (pp. [illegible]), since this is a particular source of difficulty in long reviews of research literature. All copy must be double spaced, including the references. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail.

Reprints. Fifty free reprints are given to [illegible] of articles and [illegible]. [illegible] those of early publication articles receive no [illegible].

Communications—including subscriptions, [illegible], and changes of address—should be addressed to the American Psychological Association, [illegible] Sixteenth Street N.W., Washington [illegible] D. C. [illegible] must reach the Subscription Office by the 10th of the month [illegible] month. Undelivered copies resulting from address changes will not [illegible]. Subscribers should notify the post office that they will guarantee [illegible] postage. [illegible] for undelivered copies must be made within [illegible] publication.

Annual subscription: $[illegible] (Foreign $[illegible]). [illegible] $1.50.

PUBLISHED [illegible]
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
[illegible]
and [illegible]

[illegible]

VOL. 55, NO. 5 SEPTEMBER, 1958

Psychological Bulletin

THE EXPRESSION AND REDUCTION OF HOSTILITY[1]

LEONARD BERKOWITZ
University of Wisconsin

The publication of *Frustration and Aggression* by Dollard, Doob, Miller, Mowrer, and Sears (17) in 1939 represents a milestone in the application of the methods and concepts of experimental psychology to important social problems. Perhaps stimulated by the early phases of World War II, the authors sought to translate clinical and sociological observations concerning hostile behavior into the operationally definable language of the laboratory. These men must be given credit for the attempt, whatever the merit of their hypotheses. In the tradition of science's continuing search for unifying principles, they showed that many apparently different phenomena could be "explained" employing only a relatively small number of theoretical variables and, in contrast to earlier writers, formulated their propositions in a relatively precise manner.

In addition, *Frustration and Aggression* also provides a systematic foundation for further research on aggressive behavior. The present paper is a review of studies reported after 1939 within the general theoretical framework utilized by Dollard and his associates. Inevitably, in the course of this review our attention will focus upon the question of the adequacy of this framework for the study of human hostility. Three groups of factors will be discussed: (*a*) those governing the occurrence of overt aggression, (*b*) those determining the nature and object of the aggressive act, and (*c*) those related to the reduction of the instigation to aggression.

[1] This review was supported by Research Grant M1540 from the National Institute of Mental Health, US Public Health Service.

CONDITIONS AROUSING OVERT AGGRESSION

Theoretical Principles

The fundamental hypothesis in *Frustration and Aggression* was first expressed formally in Sigmund Freud's earlier writings. He has suggested that aggression was the "primordial reaction" to the frustration occurring "whenever pleasure-seeking or pain-avoiding behavior was blocked" (17, p. 21). Acknowledging their indebtedness to Freud, the authors presented a sweeping generalization as their basic postulate: " . . . the proposition is that the occurrence of aggressive behavior always presupposes the existence of frustration and, contrariwise, that *the existence of frustration always leads to some form of aggression*"[2]

[2] A "frustration" is defined by the F-A psychologists as "an interference with the occurrence of an instigated goal-response at its proper time in the behavior sequence" (17, p. 7). This is the usage adopted in the present paper. It should be noted that other writers (e.g., 9) utilize the term to refer to the emotional state created by such an interference.

Dollard, Doob, et al. define "aggression" as

[italics mine] (**17**, p. 1). Needless to say, the second phrase drew the fire of many critics (e.g., **6**, **26**, **41**).

In a symposium on the frustration-aggression hypothesis appearing two years later, Miller (**49**) admitted that this phrase "was unclear and misleading" for two principal reasons:

> In the first place it suggests, though it by no means logically demands, that frustration has no consequences other than aggression. This suggestion seems to have been strong enough to override statements appearing later in the text which specifically rule out any such implications (cf. pp. 8–9, 19, 58, 101–102). A second objection to the assertion in question is that it fails to distinguish between instigation to aggression and the actual occurrence of aggression. Thus it omits the possibility that other responses may be dominant and inhibit the occurrence of acts of aggression. In this respect it is inconsistent with later portions of the exposition which make a distinction between the instigation to a response and the actual presence of that response . . . .
>
> Both of these unfortunate aspects of the former statement may be avoided by the following rephrasing: Frustration produces instigations to a number of different types of response, one of which is an instigation to some form of aggression (**49**, p. 338).

Dollard, et al. hypothesize that the strength of the instigation to aggression varies directly with at least three factors: "1) the strength of instigation to the frustrated response, 2) the degree of interference with the frustrated response, and 3) the number of frustrated response-sequences" (**17**, p. 28). A subsequent section adds the anticipation of punishment for aggressive behavior to the list of variables affecting the likelihood of an overtly hostile response to frustration; specific aggressive acts will be inhibited to the extent that punishment is anticipated to be a consequence of these acts (**17**, p. 33).

This list obviously is not complete; nothing is said of persistent individual characteristics, and the Frustration-Aggression group, with two slight exceptions, tends to neglect this class of factors in their psychological principles. One exception is found on p. 37. Discussing Allport's theory of trait structure, they suggest that individuals may have "generalized habits" of responding to frustrating situations with either overt or nonovert aggression. For the second exception, it is stated: "As a result of his life history any given person will carry into adult life a high or low ability to 'tolerate' frustrations and will stand at some point on a dimension of 'readiness to be aggressive' in frustrating situations" (**17**, p. 88).

Miller's 1941 article presents a further elaboration of the role possibly played by persistent response tendencies in the frustration-aggression sequence. "Instigation to aggression may occupy any one of a number of positions in the hierarchy of instigations aroused . . . [by a frustration]" (**49**, p. 338). Therefore, if, for a given individual, the instigation to aggression is lower in this hierarchy than instigations to other responses incompatible with aggression, he is not likely to demonstrate hostile behavior when confronted with a frustration situation. It is clear, then, that relatively stable individual characteristics, including the

---

any "sequence of behavior, the goal-response to which is the injury of the person toward whom it is directed" (**17**, p. 9). The authors explicitly state that they are not referring to "instrumental" aggression: "aggression is that response which follows frustration, reduces only the secondary, frustration-produced instigation, and leaves the strength of the original instigation unaffected" (**17**, p.11). Allport discusses various kinds of acts that have been labeled "aggressiveness" (**1**, pp. 355–356).

The term "instigation to aggression," as used here, denotes an aggressive drive, or more generally, a state of readiness to make aggressive responses, and *not* the frustrating event.

strength of the instigation to aggression relative to the strength of other frustration-produced instigations, should be considered among the factors affecting the likelihood of an overt hostile act.

The following section will review studies having some bearing upon the above list of factors presumably affecting the occurrence of a hostile act. Generally consistent with the theoretical expectations of Dollard and his colleagues, these results frequently add needed details to their somewhat overly general hypotheses.

Empirical Investigations

Strength of instigation to the frustrated response. An experiment and two questionnaire studies are cited as evidence for the hypothesis that the strength of the instigation to aggression is directly related to the strength of the frustrated drive. Sears and Sears measured the latency of a crying response to the withdrawal of a bottle from the mouth of a five-month old infant, with the strength of his hunger drive manipulated by varying the amount of milk that the baby was permitted to take. Assuming that crying was an aggressive reaction, their "figures indicate that as the child became more nearly satiated, i.e., as the strength of instigation decreased, frustration induced a less and less immediate aggressive response" (**17**, p. 29).[3] In the first questionnaire study, college students queried by Doob and Sears indicated that they were more likely to react aggressively in frustrating situations the stronger the drive whose goal-responses had suffered interference (**17**, p. 29). Somewhat similar results were obtained by Miller with a list of "annoyances" presented to college students (**17**, p. 30). For further studies along this latter line, an investigator might care to make use of the list of anger situations reported by Anastasi, et al. (**2**) based upon "controlled diaries" kept by 38 college women.

Getting away from the dangers inherent in the use of questionnaires, two experiments make use of the familiar Hullian principle that drive strength increases the closer the organism is to making the appropriate goal response. Seward (**73**) paired six rats in a number of different combinations after 21 hours of food deprivation. He found that the animal in possession of the food pellet, and therefore, who was closer to eating, showed reliably more aggression than the one without it. Presumably, the hunger drive whose goal responses were blocked by the behavior of the other animal was stronger in the rat possessing the food pellet. The second of these studies, carried out by Haner and Brown (**25**), provides additional support for the hypothesis of a positive relationship between closeness to the goal and drive strength. Thirty elementary school children individually played a game in which marbles had to be inserted into holes. The *E* terminated the trials at various distances from the goal of game completion, with each termination causing a buzzer to be sounded until *S* pushed a plunger to stop it. The investigators assume that the pressure exerted in pushing the plunger is a direct function of the strength of *S*'s instigation to aggression, and show that this

[3] Sears may no longer regard the child's cry as an aggressive response. In his most recent book, he distinguishes between rage or anger, on the one hand, and the desire to hurt on the other (**71**, pp. 221–222), a distinction similar to that proposed by McClelland (**44**, p. 513). The cry is an indication of rage and not aggression, as this latter response is defined by Dollard, Doob, et al.

pressure generally tends to increase the closer *S* is to task-completion. Apparently, then, the greater the strength of the drive frustrated the greater the resultant aggression—if the pressure against the plunger is indeed a hostile act and not a manifestation of something else, such as a frustration-induced tendency to exert more vigorous responses in general (**40**).

Degree of interference with the frustrated response. The second hypothesis posits a direct relationship between strength of the aggressive drive and the degree of interference with the frustrated response. The evidence utilized by Dollard, et al. in support of this principle comes from two correlational studies. First, in Miller's study of annoyances already referred to, *S*s "reported that they felt much more irritated at being completely 'off form' in their favorite sport than at being only slightly 'off form.' The critical ratio of this difference was 5.5" (**17**, p. 31). Somewhat more unsatisfactory is the second investigation by Hovland and Sears (**17**, p. 31). Reasoning that bad economic conditions should produce "a greater interference with customary goal-responses" than good business conditions, they employed the per acre value of cotton as "an index of the severity of interference with economic actions." For 14 Southern states for the years 1882 to 1930 the correlation between this index and the number of lynchings was −.67; "i.e., the number of lynchings (aggression) increased when the amount of interference increased." Mintz (**51**), however, questions the magnitude of the reported correlation. He claims the true correlation is much lower than that given by the F-A (Frustration-Aggresssion) psychologists.

Three more recent investigations can be cited as evidence for the present principle if we first assume that the degree to which some drive possessed by *S*, such as for self-enhancement, is blocked is a positive function of the strength of a hostile act directed against *S*. Explicitly making this assumption (**24**, p. 518), Graham, Charwat, Honig and Weltz presented 50 incomplete statements to 106 adolescents. The statements described five types of hostile actions that various instigators had taken against a person, with these actions ranging from a physical blow to not liking the person. The *S*s had to indicate what they thought would be the "most likely way for a person to act in such a situation." If a very hostile act directed against the individual can be considered as producing greater blocking of his goal responses than a less hostile act,[4] the present hypothesis is confirmed: "When the instigation to aggression is itself a form of aggression, the strength of the aggressive response will vary as a function of the strength of the instigation" (**24**, p. 514). In the second experiment, by McClelland and Apicella (**42**), similar results were obtained. There was a significant increase in the number of aggressive responses made in the frustrating situation when *E* increased the intensity of the derogatory remarks made to *S*. Consistent with this, French (**22**) employed 10 groups of *S*s working on three problems, and found that there was a general tendency for those individuals who received most aggression to initiate most aggression. As French notes, "Presumably those who receive aggression from others perceive these

[4] An alternative interpretation also utilized by Graham, et al. is reported below in the section, Opposing Hypotheses.

aggressors as interfering agents . . . " (**22**, p. 286).

French also compared the number of aggressive responses made during a task involving high member interdependence with the number elicited in the two low interdependence tasks. Although no quantitative data are given, French suggests that "the comments of both the observers and the subjects prove that a member suffers more interference by the others" on the former problem. He reports more aggression per minute directed against other members of the group in this problem than in the low interdependence. A statistical analysis is given (**22**, pp. 286–287), however, that adds further corroboration to the present hypothesis.

Number of frustrated response-sequences. The third hypothesis maintains that the total strength of the instigation to aggression is a positive function of the number of frustrated response-sequences. The only evidence for this principle given in *Frustration and Aggression* is of an anecdotal nature. Thus, in an experiment by Sears, Hovland and Miller, it was observed that "a man who had previously been a willing subject for several arduous experiments complained vigorously at having to give free associations to fifty stimulus words" (**17**, p. 32). Apparently the minor frustrations he had suffered had combined to produce a hostile response of greater strength than would ordinarily be expected from the immediate frustrating situation alone. Incidental support for the present hypothesis is found in a more recent experiment by Thibaut and Coules (**79**). In one part of the investigation two conditions were established after all *Ss* had been angered by notes supposedly written by a fellow college student. In one of these conditions *Ss* were interrupted by *E* after the instigation for about three minutes, while the remaining *Ss* were allowed to write one note to the instigator before this informal interruption occurred. Then, *Ss* in both treatments continued communication with the instigator. There was a significantly greater volume of aggression communicated to their fellow student by the former *Ss*, even though the experimenter rather than this person was responsible for the interruption.

Several points may be noted here if we assume, as seems likely, that there was little cathartic reduction of hostility in the writing of the single note. First, the interference with the aggressive responses apparently was in itself frustrating, as Dollard, et al. have hypothesized (**17**, p. 40), and this seems to have led to increased hostility. Second, since the insulting notes probably had made all *Ss* angry (in the absence of a control group in the present case it is impossible to demonstrate this unequivocally), the hostility produced by the interference to the aggression presumably had added to the hostility created by the faked notes.

Miller's 1941 paper (**49**) contains an hypothesis that is closely related to the present problem. He proposes that frustration-induced instigations to responses other than aggression weaken as the frustration persists. Since the earlier statement maintains that the aggressive effects of a number of interferences summate, both hypotheses yield the same prediction: there is a greater probability of overt aggressive acts with repeated frustrations.

Otis and McCandless (**56**) attempted to test Miller's newer hypothesis in an experiment with 63 preschool children placed in an eight-

trial frustration task. The *Ss* showed significant increases in "aggressive-dominant" behavior scores from the first four to the last four frustration trials, and showed reliable decreases in a class of behavior found to be incompatible with aggression, "submissive-complaisant" behavior. In contrast to this support for the present hypothesis, McClelland and Apicella (**42**) obtained negative results in an experiment employing college students on only two frustration trials.

Inhibition of aggression. Basically, according to the F-A group, "the strength of the inhibition of any act of aggression varies positively with the amount of punishment anticipated to be a consequence of that act" (**17**, p. 33). The evidence presented for this proposition stems from the above mentioned questionnaire study by Doob and Sears (**17**, p. 35). When asked to indicate how they had responded to frustrating situations they had actually experienced, the 185 college students were more likely to report an overtly hostile act when the satisfaction obtained from making this response was greater than the anticipated punishment, while the smallest frequency of overt aggression occurred when the anticipated punishment was greater than the anticipated satisfaction.

On the basis of records of actual experiences kept by 120 adults, McKellar (**47**) reports that acts of aggression tend to be directed against an object from which retaliation is unlikely. The most direct recent support for the hypothesis dealing with the effects of punishment comes from an investigation conducted by Chasdi and Lawrence (**10**). Twenty-three preschool children were divided into two groups: 12 in an experimentally punished group and the remainder in a control group. Both groups had four sessions of doll play with the experimental group receiving punishment for the expression of aggression in Session II. There were significant differences between the two groups on both the frequency and intensity of aggression in Session III. The difference was still in the predicted direction during Session IV, but it is no longer significant. According to the experimenters, this latter change "is consistent with a session to session increase in aggression resulting from a decrease in aggression anxiety due to the reinstatement of permissiveness" (**10**, p. 522).

In this study, as well as others employing doll play measures of hostile tendencies (e.g., **4**, **37**, **84**), there was a significant increase in aggressive manifestations from the first to the last sessions for all groups, presumably because the children learned they would not be punished for displaying hostility in the play situation. Somewhat in accord with this last hypothesis, Feshbach (**19**) found that boys initially low in aggressiveness tended to increase in their display of overt hostility following a series of permissive free play experiences. This increase was not obtained with the girl *Ss*.

Seward's studies of aggressive behavior in rats also illustrate the inhibitory effects of punishment. In one experiment (**74**), he observed that the day after a fight the average loser made fewer aggressions or advances than before not only against the winner but also against other rats. Seward attempts to show how stable dominance-submission relationships can develop through the conditioning of the fear drive in one animal to the presence of another (**72**).

Something analogous to this probably is involved in the class of dominance-subordination relationships we term social status. For one thing, human beings undoubtedly learn they will be punished by the dominant, high status figures in their social groupings if they aggress against them. Thus, it is not surprising to see *Ss* express less hostility toward frustrators vested with power and authority by society than toward lower status instigators (**13**, **24**, **33**, **80**).

Social groups may not only teach the individual to anticipate punishment for disapproved actions, they also may influence the reaction to these anticipations. One of the first studies to demonstrate this was carried out by Wright (**82**). Pairs of nursery school children, grouped in terms of whether they were "strong" or "weak" friends, were frustrated by the adult *E*. The pairs of strong friends were more likely to express aggression, including direct physical attacks, against the experimenter than the pairs of weak friends. An experiment by Pepitone and Reichling (**58**) suggests that the variable probably largely accounting for this type of effect is the liking of the group members for each other. After they had been treated in an unjust and arbitrary manner by an instigator who then left the room, male college student members of experimentally-produced high liking (high cohesive) groups expressed significantly more hostility toward the instigator than did the members of the low cohesive groups. The former *Ss* apparently felt fewer restraints in attacking *E* and the experiment.

Conceivably, the lowering of restraints accompanying high liking among the group members also can result in a greater volume of intragroup overt hostility. French, in the experiment described earlier (**22**), studied two types of groups working under conditions in which the members were likely to see each other as frustrators. In one type of group the members had known each other for some time, while the *Ss* in the remaining groups did not have this history of previous contact. The observers recorded significantly more overt aggression directed toward other group members in the former groups, supposedly because the higher liking resulting from the greater degree of acquaintance had lowered the members' inhibitions.

Levin and Turgeon (**37**) also have shown that the presence of a familiar person is associated with a relatively great release of overt aggression, but the reason for this finding in their experiment is not clear. They found that aggressive doll play behavior significantly increased when a child's mother was present, and tended to decrease (nonsignificantly) when an adult female stranger watched the child's play. The reduction in hostile behavior in the presence of the stranger cannot be traced to a general inhibition of activity in this condition, and the authors speculate that the decrease is "a reflection of the child's training in 'company manners,' to behave well in front of strangers." In opposition to the hypothesis that the mother's presence served to weaken the inhibitions against the expression of aggression, there are suggestive, but nonsignificant, indications that the mother's presence actually instigated such actions, particularly in the case of the boys.

As will be discussed more fully later, the F-A psychologists propose that the inhibition of direct acts of aggression increases the likelihood of

either: (*a*) indirect acts of aggression against the instigator, (only the form of the aggression is changed, not the target) or (*b*) displaced aggression in which a target other than the frustrator is attacked. A third inhibitory reaction to internal feelings of anger also has been described (**55**): withdrawal from communication with the instigator. Although the withdrawal, in its effect, may be an indirect attack, as in the case of "snubbing," it is likely that the break in communication is motivated by: (*a*) inhibition of direct hostile attacks against the instigator and (*b*) the unpleasant effects within the individual created by the frustration of the instigation to aggression. Thibaut and Coules (**79**) have shown that interference with the occurrence of aggressive responses can become a source of discomfort. Whatever the reason for the withdrawal reaction, Newcomb (**55**) points out that hostile attitudes often persist because they lead to a breakdown in communication with the object of the hostility. Thibaut and Coules also report findings partially supporting this hypothesis. On the basis of each *S*'s initial attitudes toward his assigned co-worker, the sample was divided into those initially hostile and those initially friendly toward their partners. (These were attitudes that the *S*s had brought into the experiment.) The initially hostile *S*s produced a significantly smaller volume of communications to their partners prior to the experimental instigation to aggression.

Characteristic individual differences. There can be little doubt that individuals differ reliably in the readiness with which they display hostile behavior. Documentation for this point, if it is needed, can be found, for example, in a study conducted by Yarrow (**84**). Based on a sample of 60 preschool children, indices of aggression computed from the first doll play session were found to be significantly related to the indices obtained in the second session following a frustration. The Pearson product-moment *r*'s ranged from .38 to .77 for the entire group. Instead of demonstrating that these individual differences exist, therefore, the more important research problem is to isolate the classes of relatively stable characteristics that affect the probability of an aggressive reaction to a given frustration.

One factor warranting further investigation involves apparent differences in the characteristic strength of the instigation to aggression. Research findings in this area point to a number of complexities that may have to be unraveled before personality tests can be used to predict the probability of hostile behavior at a satisfactory level of confidence. For example, there is an apparently paradoxical result of the Thibaut and Coules experiment. After receiving the insulting notes, the initially hostile *S*s communicated a significantly *smaller* amount of aggression to the supposed insulter than the initially friendly *S*s. A similar difference, obtained in another experiment, suggests that this may be a reliable pattern. Hokanson and Gordon (**27**) classified their *S*s, 40 male college students, as either high or low hostility expressers on the basis of their responses to a scale of "manifest hostility." The *S*s scoring at the extreme ends of this scale were placed in either a situation designed to arouse relatively strong hostility, or in a low arousal situation, and then allowed to express aggression in fantasy (to TAT pictures), and in overt behavior. Relative to the comparable *S*s in the low arousal condition, there

was a significant tendency for the "strongly aroused" *Ss* with low manifest hostility scores to express more hostility on the TAT, while the similarly "aroused" high manifest hostility *Ss* gave fewer aggressive responses. No significant relationships were found in the overt behavior situation.

There is no unequivocal explanation for the above findings. It may be, as Hokanson and Gordon assumed, and as Feshbach suggests (**19**, p. 452), that the initially "friendlier" *Ss* in these studies were low in overt aggressiveness more because of high inhibition in the original (or test-taking) situation than because of low drive strength. As was mentioned earlier, Feshbach (**19**) obtained results somewhat in accord with this interpretation. Similarly, Bach (**4**) found a tendency, though nonsignificant, for 12 preschool children previously rated by their teachers as showing "little destructive aggression" to have higher doll play aggression in the last of four permissive play sessions than the 14 children previously rated characteristically high in aggression. According to this view, then, the permissive experimental situation lowered the inhibitions felt by the *Ss* ostensibly low in characteristic aggression. As a result, their stronger instigation to aggression led to stronger aggressive responses.

Another possibility is suggested by Miller's "approach-avoidance" conflict model (**50**). The *Ss* with high initial hostility scores may have had a truly stronger aggressive drive, or perhaps a stronger readiness to respond aggressively to an instigation, than the *Ss* with low scores. The effect of the hostility instigation, then, was to produce greater anger in the former *Ss*. However, as a consequence of this hostility arousal, the *Ss* also were likely to have stronger guilt feelings, or aggression-anxiety, when they became aware of their hostility. In reaction to this emotion, these *Ss* then suppressed the overt manifestations of aggression. Instead of stating that anxiety inhibits aggression only prior to the experimental situation, as the first hypothesis maintains, this latter hypothesis suggests that it inhibits the expression of aggression during the experiment.

There is some support for this alternative explanation in an experiment by Clark (**12**) dealing with another drive that also is often socially disapproved, sex. Clark found that male college students, whose sex drive presumably was heightened as a result of seeing pictures of nude women, expressed significantly lower "manifest sex" to TAT cards than the *Ss* in the nonaroused group (*Ss* not shown these pictures and therefore presumably having a low drive level). It was not until the *Ss* were placed in a highly permissive situation (with their superegos "dissolved in alcohol") that the sexually-aroused group had higher manifest sex scores. Thus, when the arousal of the socially disapproved drive was likely to produce guilt feelings, the aroused *Ss* overtly displayed a lower level of this drive than the nonarousal *Ss*.

Interestingly enough, Clark also found that the sexually-aroused *Ss* under nonalcohol conditions gave significantly fewer aggressive responses to the TAT cards than the control *Ss*. Evidence is presented suggesting that the former group "was not only anxious about expressing manifest sex, but also was anxious about expressing manifest aggression . . . " (**12**, p. 56).

The present data do not permit us

to choose between the above alternatives, and indeed, both may be correct. The first possibility may account for the increase in overt aggression for *Ss* characteristically low in this mode of behavior, the second for the decrease shown by the supposedly "hostile" *Ss*. But whatever the explanation, the results obtained by Thibaut and Coules and Hokanson and Gordon add further documentation to a point made by many writers (e.g., **17**): the amount of overt aggression typically displayed by an individual is not necessarily a direct function of the strength of the instigation to aggression within him.

Two hypotheses may be formulated which relate characteristic overt hostility to the strength of the individual's instigation to aggression by extending the F-A propositions regarding anticipation of punishment. The first hypothesis maintains that an individual's characteristic level of overt aggression is more likely to coincide with the strength of his aggressive drive the more his important socializing agents (e.g., parents) have permitted or encouraged aggressive behavior.

There is support for this hypothesis in two recent investigations. Lesser (**35**) obtained a significant +.43 product-moment correlation between fantasy aggression scores and reputation among peers for overt aggression for 23 elementary school boys whose mothers, when interviewed, indicated they were relatively supportive of aggression. The correlation was negative (−.41) in the case of 21 boys whose mothers said they discouraged aggression. Similarly, contending that lower-class culture often encourages aggression, Mussen and Naylor (**54**) found a significantly positive association between ratings of overt aggression and number of aggressive TAT themes for a sample of 24 lower-class juvenile delinquent boys. In both of these studies, therefore, the results suggest that the level of the individual's fantasy aggression is likely to approach the level of his overt aggression when important figures who can either reward or punish his hostile behavior tend to reinforce this type of response. Although investigations mentioned earlier (**10, 12, 27, 37**) indicate that fantasy aggression is not always directly proportional to the strength of the instigation to aggression, this direct relationship is plausible in the case of the present two studies. The agents primarily involved in the socialization of the Lesser and Mussen and Naylor *Ss* presumably punished aggressive behavior, including the kind of indirect hostility involved in the fantasy tests, relatively infrequently. Sears, Maccoby and Levin also show that characteristic overt aggression is associated positively with the socializing agents' permissiveness for aggression (**71**, pp. 258–259).

The second hypothesis, relating characteristic hostile behavior to the level of the aggressive drive, states: Frequent punishment for the expression of aggression increases the strength of the instigation to aggression. This obviously is consistent with the F-A principle, previously mentioned, that "interference with direct aggression constitutes in itself an additional frustration" (**17**, p. 40). A variation of this hypothesis, first postulated by Whiting, is employed by Sears and his colleagues. They argue that frequent punishment for hostile behavior produces aggression anxiety which interferes with the occurrence of aggressive acts. Therefore, the greater the aggression anxiety, the greater the frustration, and consequently, the stronger the re-

sulting instigation to aggression (**70**, p. 218).

Two recent studies conducted by Sears and his associates have results generally consistent with this reasoning. In the first of these (**70**), the mothers' statements of how punitive they were with regard to aggression correlated +.50 with teachers' ratings of the aggressiveness of 21 nursery school boys, and +.60 with the observed frequency of this behavior in 16 15-minute nursery school time samples. These correlations were in the opposite direction for the 19 girls in the group, −.04 and −.41 respectively. To account for the sex difference, the investigators suggest that girls are more likely to identify with their mothers than boys, and therefore, a given amount of punishment by the mother is felt more strongly by the girls. As a consequence of the strong aggression anxiety, the highly punished girls exhibit a stronger tendency to inhibit direct overt aggression (**70**, p. 219). When the children's activity level was taken into consideration, "the more severely punished girls were relatively [but not reliably] more aggressive than the less severely punished" (**70**, p. 216).

The second study, reported more recently by Sears, Maccoby and Levin, does not show sex differences in reactions to maternal punitiveness. For both boys and girls, mothers who indicated that they punished their children's aggressiveness severely tended to describe their children as highly aggressive (**71**, p. 259). While admitting the possibility that maternal punishment was a response to the child's aggressiveness, the authors argue that this punishment also is likely to have bred counter-aggression in the children.

To summarize, the extent to which important socializing agents either reinforce or punish the individual's aggressive behavior appear to be important determinants of the degree to which the level of his overt hostility coincides with the strength of his aggressive drive. Evidence indicates that there is a positive association between these two variables if the socializing agents permit or encourage the expression of hostile behavior. Punishment of this behavior, on the other hand, seems to increase the strength of the instigation to aggression, although, particularly if it is intense enough, inhibiting direct overt aggression.

Past learning experiences apparently can also modify the relative dominance of the instigation to aggression in the hierarchy of frustration-produced instigations by reinforcing other modes of response. Experiments suggest that it is possible to teach nursery school (**32**) and elementary school (**14**) children to react "constructively" to frustrating situations.

Several writers, when describing characteristic modes of behavior, refer to constructive frustration reactions as a manifestation of a high "frustration tolerance" (e.g., **61**, p. 385), while others (e.g., **7**) prefer to employ the construct, "strong ego control." In the last-mentioned paper, Block and Martin show, in an extension of the classic Barker, Dembo and Lewin (**5**) study, that "ego-control capacity," defined independently in terms of ability to delay gratification and low cosatiation,[5] is predictive of individual differences in reactions to frustration. In a sample of 22 preschool children, those with low "ego-control" responded to a frustration with less constructive

[5] Cosatiation is the satiation of one drive resulting from the satiation of another drive.

play behavior and acted aggressively more frequently than the other children. Consistent with this finding, Livson and Mussen (**39**) report that in a sample of 36 nursery school children the above two measures of ego control were negatively related to indices of aggressive behavior obtained in a two-week observation period (the correlation was significant only for the cosatiation measure).

Other individual characteristics apparently affecting reactions to frustration have been described as well. Rosenzweig found that *S*s are most likely to react aggressively to frustration experiences when they possess relatively great confidence and ability (**60**). According to Zander (**85**) extroverted *S*s tend to show more aggression against the frustrating experimenter than introverted *S*s. Lindzey indicates that *S*s high in minority group prejudice are more likely to react emotionally to frustrations but not necessarily with aggression (**38**). The sex of the *S* has frequently been found to be a determinant of response to frustration, with males usually reported as displaying more aggression than females (**65**; **71**, p. 264; **72**; **84**). Social class membership often is said to be related to the extent to which aggression is a characteristic frustration-reaction (e.g., **46**, **64**), but this has not always been confirmed (e.g., **71**, p. 265).

Finally, there is suggestive evidence concerning the home environments that affect the learning of persistently hostile behavior. Allinsmith found that middle-class (but not lower-class) males are most likely to feel guilty about their aggressive death wishes when their mothers have employed psychological rather than physical discipline (**48**, p. 155). Sears, Maccoby and Levin (**71**) show that mothers who characteristically employ physical punishment and/or who are relatively cold in their affectional interaction with their children also tend to describe their children as highly aggressive. (The *r*'s are very low but still significant.) Bornston and Coleman (**8**) have somewhat similar results with regard to maternal coldness. After administering the Rosenzweig P-F Study to 56 college students and a parent attitude survey to the students' mothers, they found that the frequency of aggressive (extrapunitive plus intrapunitive) responses on the personality test was significantly positively related to both maternal domineering and ignoring attitudes toward the child. The Sears group also reports that the prolonged absence of the father from the home produces a relatively low level of fantasy aggression in preschool boys, but not girls (**65**, **69**). Other studies along these lines are cited by Child (**11**, pp. 669–672). In conclusion, both Sears, Maccoby and Levin (**71**, p. 266) and Child (**11**, p. 670) offer suggestions as to how parents should respond to a child's aggressiveness if they want to avoid producing a characteristically hostile child.

Opposing Hypotheses

Most of the attacks upon the *Frustration and Aggression* hypotheses concentrate on the basic postulate. As was indicated earlier, many writers have disputed the contention that frustration always leads to some form of aggression, and their objections are not necessarily overcome by Miller's 1941 revision (**49**). This later statement still proposes that an instigation to some form of aggression is an invariant product of frustration, even though other instigations may be stronger, while the critics usually suggest that only cer-

tain classes of frustrations give rise to hostility. Substantially in agreement with one another, they maintain that aggression is a reaction only to threatening frustrations or attacks, and does not occur in response to sheer deprivations (e.g., **41**, **61**).

However, these critics have not satisfactorily overcome the major difficulty inherent in this alternative formulation: the a priori specification of the operations distinguishing "threat" or "attack" from "deprivation." It obviously is not enough to state that the former involves the security of the organism while the latter does not; for example, even minor deprivations can provoke aggression at times (**17**, p. 31), suggesting either that this alternative hypothesis is incorrect, or that deprivations can become threats depending upon changing conditions of the organism. Those arguing for the latter point of view can easily find themselves holding a circular definition of "threat." Taking a somewhat different approach, McClelland hypothesizes that an individual is most likely to display defensive behavior, including aggression, when he is faced by an unknown danger and has no problem-solving response readily available (**44**, pp. 504–505).

Clearly, as Sears has pointed out in the 1941 symposium on the *Frustration-Aggression* hypotheses, one of the major problems in the study of hostile behavior has to do with "the determination of the specific factors which cause one kind of frustration-reaction rather than another to occur" (**66**, p. 345), and the question of whether qualitatively different classes of frustrations produce different reactions must be included as one of the more important aspects of this problem. It also is clear that an unequivocal answer to this question will come only from carefully controlled experiments rather than from questionnaire investigations posing hypothetical situations to which *Ss* are asked to respond.

Three papers appearing in the last several years illustrate the difficulties in obtaining answers to this question. Graham, et al. (**24**) state that an aggressive act directed against *S* can be interpreted not only as an interference with some drive such as self-esteem, (this is the interpretation employed in our earlier citation of this study) but also as an attack (or threat). In order to demonstrate that only the latter produce aggressive reactions, it is necessary to establish independent operational definitions of "degree-of-blocking" and "strength of attack."

While such a distinction conceivably could be made in a questionnaire investigation, this type of procedure is particularly susceptible to the intrusion of *Ss*' own hypotheses. They may be more likely to respond in terms of what they *believe* is, or should be, the case when presented with questionnaire descriptions of hypothetical situations than when actually confronted with a frustrating event. Pastore (**57**), for example, has argued that the frustrating situations referred to by Doob and Sears in their questionnaire study (**17**, p. 29), evoked hostile reactions primarily because these frustrations were unreasonable or arbitrary in nature. Employing a similar procedure, he showed that *Ss* are significantly more likely to indicate they would react aggressively when given hypothetical arbitrary frustrations than when the frustrations appear "reasonable." (Cohen (**13**) has obtained identical results with the "arbitrary—reasonable frustration" variable utilizing the same procedure.) However,

Pastore acknowledges that the *S*s could have inhibited aggressive responses when the frustrations were not arbitrary because hostility was "unreasonable." If this is the case, we might also speculate that this inhibition is more evident in a questionnaire investigation than in an actual experiment.

Pastore also proposes another explanation of his results that is similar to an hypothesis advanced by several writers (e.g., **85**). He suggests that the arbitrary frustration involves the frustration of an expectancy (e.g., an *S* had expected to get on a bus. When it suddenly passed him by, this expectancy was frustrated). According to this view, then, aggressive reactions are most likely when the expected attainment of a goal is blocked. Bateson presents an essentially similar hypothesis in his discussion of cultural factors affecting the frustration-aggression relationship (**6**). Because of continual frustration, particularly at the hands of their mothers, the Balinese presumably do not learn to expect strong satisfaction, and therefore, Bateson argues, "they are infinitely willing to suffer interruption" (**6**, p. 353).

The Nature and Target of the Aggressive Response

Theoretical Principles

The F-A discussion of factors affecting the nature and target of the aggressive response, particularly as elaborated by later theoretical papers (e.g., **50**), must be considered among the prime examples of the benefits to be derived from the wedding of learning theory and psychoanalysis. The concept of generalization, developed largely through laboratory investigations of learning, is integrated in a highly fruitful manner with the often rich insights provided by psychoanalysis.

The first specific hypothesis listed by Dollard and his associates has to do with the strength of various hostile tendencies. "The strongest instigation, aroused by a frustration, is to acts of aggression directed against the agent perceived to be the source of the frustration and progressively weaker instigations are aroused to progressively less direct acts of aggression" (**17**, p. 39). The second half of this hypothesis appears to be based upon the notion of response generalization; "the more direct acts of aggression will be those which are more similar . . . to the act of most direct aggression" (**17**, p. 39). Indirect evidence supporting the first part of the hypothesis is reported in the previously cited study by Doob and Sears. Their *S*s indicated that acts of direct aggression were more satisfying to them than other, more indirect, forms of aggression.

In an interesting extension of the principle that aggression is directed against the object perceived to be the source of the frustration, Dollard, Doob, et al. hypothesize that self-aggression results when the self is seen as the frustrating agent. Therefore, "self-restraint of an act of aggression should instigate aggression against the self" (**17**, p. 48).

Since the interference with a directly aggressive response is an additional frustration adding to the total strength of the instigation to aggression (**17**, p. 40), as has been shown earlier, another hypothesis is derived: "The greater the degree of inhibition specific to a more direct act of aggression, the more probable will be the occurrence of less direct acts of aggression" (**17**, p. 40).

Similarly, if acts of aggression directed against a given object are prevented there will be a tendency for the individual to attack other objects. In a later paper (**50**), Miller

shows that this phenomenon, labeled "displacement" by psychoanalytic theory, can be understood as a case of stimulus generalization. He also points out the importance of determining whether the direct act of aggression is prevented by the absence of the perceived instigating object or by conflict (produced by anticipation of punishment for aggression or by aggression-anxiety). If it is the former, displaced responses will occur to other similar objects and the strongest hostile response will occur to the most similar object present. If there is conflict, however, on the assumption that the gradient of generalization of the interfering responses is steeper than that of the aggressive responses they inhibit, it is predicted that the strongest displaced hostile act will occur to stimulus objects which have an intermediate degree of similarity to the original object. Increasing the relative strength of the inhibitory response serves to shift the displacement to a less similar target, while a more similar target will be attacked the greater the relative strength of the instigation to aggression. Increasing the aggressive drive also raises the likelihood that increasingly dissimilar objects will be capable of evoking responses.

Three studies are reported by the authors of *Frustration and Aggression* in support of the principle of displacement. First, it was shown that rats who had been trained to strike at another rat would attack a celluloid doll when no other rat was present (**17**, p. 42). Second, in a study carried out by Miller and Bugelski, experimentally frustrated *Ss* tended to rate their friends (who had not been responsible for the frustration) lower on a personality scale than did control *Ss* who had not been subjected to these frustrations (**17**, pp. 42–43). The last experiment, also conducted by Miller and Bugelski, is frequently cited by proponents of the "scapegoat" theory of prejudice. Boys in a CCC camp displayed less favorable attitudes toward Mexicans and Japanese after a frustration (**17**, pp. 43–44).

Empirical Investigations

French's previously mentioned study of "organized" and "unorganized" groups contains observations supporting the view that aggression tends to be directed against the agent perceived to be the source of the frustration (**22**, p. 286). Pepitone and Reichling (**59**) also report corroborating observations: over all groups in their experiment the major target for hostility (in terms of mean seconds of expression) was the insulting *E*, rather than Psychology, the experiment, or the physical setting.

McClelland presents suggestive evidence for the proposition that self-restraint of an act of aggression leads to self-aggression (**44**, p. 516). Utilizing the results of a study by MacKinnon, he argues that intropunitives (those who tend to blame themselves for a frustration, i.e. who are high in self-aggression) had been restrained from aggression and other antisocial acts as children by the development of internal standards of conscience. "Consequently these internal restraints would more often be perceived as frustrators by them and aggression should consequently be more often directed against the self for imposing them . . . ." On the basis of a significant negative correlation between intensity of overt aggression and self-aggression, Thibaut and Riecken (**80**) consider acts of aggression toward the self to be a "sensitive symptom of strength of inhibition of aggression."

French (**22**) also reports findings in support of the hypothesis that indi-

rect acts of aggression are more likely to result the greater the inhibition of direct aggression. The "organized groups" in his study exhibited less social restraint than the "unorganized" groups, presumably because the members of the former groups had known each other longer. Thus, where there were 61 instances of direct aggression (physical attacks, verbal hostility) in the organized groups, the inhibiting social restraints in the unorganized groups apparently produced only indirect acts of hostility (blaming, dominating others, arguing, hostile jokes). Explicitly employing the concept of response generalization, Dinwiddie (**15**) showed that *S*s with high scores on a scale of "social anxiety" exhibit reliably more indirect aggression responses to the Rosenzweig P-F Study than the *S*s low on this anxiety measure with hostility level (as assessed by a manifest hostility scale) held constant statistically. The "social anxiety" is assumed to serve as an aggression-inhibitor.

In contrast to the previously mentioned experiments demonstrating a shift in the target of aggression (i.e., displacement) when direct attacks are prevented, several studies have yielded negative results (e.g., **21, 38, 76**). The lack of displaced aggression in two of these investigations (**38, 76**) is regarded by some writers as limiting the generality of the displaced hostility—"scapegoat" theory of prejudice. Thus, Allport comments, "It is not true that a defenseless minority is always chosen for displacement purposes" (**1**, p. 351).[6]

However, Miller's conflict model (**50**) proposes that displacement should occur only in situations resembling but yet psychologically distant (in terms of either time, space or similarity) from the situation originally eliciting the inhibited aggressive responses. A given stimulus object will not be attacked solely because it is available if it is either too close to or too far removed from the original instigating situation on some appropriate stimulus-generalization gradient. Doll play investigations of hostility can provide relevant data. Since fantasy activities in a nursery school setting probably are sufficiently removed from the instigating home environment, aggressive tendencies that are punished in the home situation, and therefore, that are inhibited, should become manifest in doll play. Several studies, employing doll play measures of aggression, demonstrate that children who are severely punished at home for aggressiveness exhibit greater fantasy aggression than less severely punished children (**10, 36, 70,** also discussed in **67, 68**).

Two recent correlational investigations also support Miller's analysis of displacement. Murney (**53**) asked judges to rank 20 figures taken from Schneidman's Make-A-Picture-Story Test on the basis of a "global impression" of their similarity to the figure of an army officer also taken from this test. An agreed-upon similarity continuum was constructed employing 10 of the figures. The *E* then told a story to each of 90 male patients at a VA center in which an army officer arbitrarily frustrated a private, and they were asked to describe what the private would do. They were encouraged to give an aggressive response and could choose any of the 11 figures (including the officer) as the object of aggression. Aggression drive and anxiety scores for each *S* were derived from TAT responses. The

[6] Feshbach and Singer (**20**) recently have obtained results consistent with the displaced hostility—"scapegoat" theory of prejudice.

stronger *S*'s instigation to aggression the more similar the figure chosen as the aggressive target was to the instigator ($r = .82$ with anxiety held constant). The object for the target of displaced hostility shifted in the direction of greater dissimilarity with increased anxiety ($r = -.72$ with the aggressive drive measure partialled out).

Wright (**81**) analyzed 12 folk tales from each of 33 societies for signs of displaced aggression and related these to a measure of aggression anxiety based upon ratings of the extent to which children were punished for aggression. The greater the aggression anxiety the less similar both the object and the agent of the aggression was to the hero of the folk tale. Wright also found that the intensity of the fantasy aggression increased reliably with the severity of the punishment for aggression.

Catharsis

Theoretical Principles

As described by Dollard and his associates, there are many points of similarity between obviously physiological drives, such as hunger and sex, and the instigation to aggression. This assumed similarity is nowhere seen more clearly than in their discussion of catharsis. When an organism makes an appropriate goal response, such as eating, the strength of the relevant physiological drive (in this case, hunger) is reduced. Causing injury to another, particularly the frustrating agent, is held to be the aggressive drive's goal response, and therefore, the occurrence of an act of aggression supposedly reduces the instigation to aggression (**17**, p. 50).

Their discussion of catharsis also includes a reference to an important implication of the joint operation of catharsis and displacement: "With the level of frustration held roughly constant, there should be an inverse relationship between the occurrence of different forms of aggression" (**17**, p. 51). Just as displacement to other forms of aggression follows the inhibition of any aggressive response, the cathartic effect resulting from the expression of a hostile act should reduce the instigation to other aggressive behaviors.

Two admittedly "slender threads of evidence" are presented for this derivation (**17**, p. 52). First, in a sleep-deprivation experiment involving six male college students, the *S* who expressed the most overt aggression injured himself least in a self-administered algesimeter test. Conversely, the *S* rated as expressing the least overt aggression inflicted the most injurious pressure upon himself. Second, in an experiment by Miller and Bugelski, *S*s who gave their frustrating partners relatively low ratings "did not drop so markedly in their own ratings of themselves, the correlation between the ratings of partner and self being $-.3$ (S.E. $= \pm .1$)." In both experiments, therefore, aggression directed against others presumably reduced the intensity of self-directed aggression, and/or vice versa.

It is important to note that this chapter also suggests two limiting conditions (unfortunately relegated to a footnote on p. 50); (*a*) the cathartic reduction of hostility may only be temporary if the original frustration persists, and (*b*) "repetition of a mode of release may presumably produce learning of it." A third condition possibly limiting the generality of the catharsis hypothesis is pointed to by Morlan (**52**). Dollard earlier (**16**) had shown how an aggressive act committed by an individual might

stimulate him to further aggression. Aggressive responses (either overt or covert) can provoke a fear of retaliation which, in turn, increases the instigation to aggression (presumably because the fear of retaliation, in itself, is a frustration). In summary, according to the F-A group, the occurrence of an aggressive act reduces the instigation to aggression, unless (*a*) the frustration persists, adding further strength to the aggressive drive, (*b*) the aggressive behavior becomes a learned response, or (*c*) implicit verbal responses are aroused or aggression-anxiety is produced which lead to further frustrations.

The following section is a review of studies having relevance to the catharsis hypothesis. As will be shown, the results tend to be equivocal for two main reasons. They either fail to take cognizance of the above mentioned limiting conditions (e.g., not controlling the frustrations suffered by the *S*s), or they fail to distinguish between aggressive responses and the instigation to aggression. Obviously, this latter failure means the investigator cannot demonstrate clearly that the decrease in hostile behavior is due to drive reduction and not to response inhibition.

Empirical Investigations

The studies can be organized in terms of the types of situations in which *S*s express aggression. Several investigations attempt to assess the strength of the instigation to aggression in children following a series of experiences in which aggressive behavior is permitted or encouraged (**3, 4, 19, 34**). Some measure the drive strength in young adults after athletic contests (**29, 30, 77**), while others deal with experimentally aroused hostility and then seek to determine the level of the residual aggression after the occurrence of aggressive responses (**18, 22, 58, 74, 78, 79**).

To discuss these in order, two studies deserve special comment because of the widespread interest in play therapy as a psychotherapeutic technique. Kenny (**34**) provided an experimental group of 15 first-grade children with two "catharsis" situations utilizing a play therapy technique. A matched group of 15 children in the control condition spent an equal time playing on the swings or with a jig-saw puzzle. Using the first five episodes of the Korner Incomplete Story Test for the assessment of initial aggressive drive strength, and the last five episodes of this test for the final hostility measure, it was found that the *control* group showed a decrease in total aggression scores which was significantly greater than the slight decrease in the experimental group. Feshbach (**19**) also found no evidence of a cathartic reduction of hostility as a function of the usual play therapy. As was mentioned earlier, he observed that boys initially low in aggressive behavior increased significantly in overt aggression after a series of permissive free play experiences, contrary to the tendency noted by Kenny. There was no significant effect for the girl *S*s. Feshbach points out that the *S*s in this experiment were not protected from continuing sources of frustration so that the catharsis hypothesis cannot be discarded. However, he also found that play with aggressive toys seemed to promote more inappropriate later aggression than play with neutral toys.

While these investigations appear to threaten several important assumptions involved in the use of play experiences as a therapeutic tech-

nique, they do not seriously attack the F-A catharsis hypothesis. There are a number of obvious flaws in their design: (*a*) the *Ss* in these studies were not isolated from continuing frustration; (*b*) the amount of frustration provided by the experimental tasks was not controlled; (Kenny's *Ss* given the "cathartic play" situation may have been more interested in the swings or jigsaw puzzles so that they actually were angered more than the control *Ss*); and (*c*) there is no information as to the amount of aggressive behavior expressed in the various conditions prior to the administration of the residual hostility instrument; (thus, the *Ss* playing on the swings may have made many more aggressive responses than the experimental *Ss*).

Appel (**3**), obtained results consistent with the catharsis hypothesis, but alternative explanations are available in this case as well. Children who had been in a nursery school that permitted fighting showed less hostility later in kindergarten than children coming from a nursery school that discouraged fighting. However, as Morlan notes, the former children may have fought less because their greater experience with quarrels had led them to develop better techniques of dealing with one another (**52**, pp. 3–4).

The studies of individuals engaged in combative sport illustrate the necessity of distinguishing between aggressive drive strength and response inhibition. Johnson and Hutton (**30**) administered the H-T-P test to eight collegiate wrestlers under three conditions: (*a*) before the wrestling season; (*b*) four to five hours before the first intercollegiate match of the season; and (*c*) the morning after the match. They report that aggressive feelings increased prior to the wrestling match, supporting Dollard's self-stimulation hypothesis, and then decreased considerably ("in most cases to a level below condition A") following the match. However, the reliability of these differences is not known since no statistical data are presented. Nor is it stated whether the psychologist who scored the protocols knew the conditions under which they were obtained. This type of research obviously requires some control over the intrusion of the projective test scorer's own hypotheses. In another inadequately reported study (although some statistical data are given), Husman (**29**) obtained the responses of collegiate boxers, wrestlers, cross-country runners and a group of nonathletes to three projective tests, including the TAT. Statistically significant differences were found among the various groups employing the TAT measures of aggression, with the boxers showing considerably less aggression than the others. However, the boxers also were rated on the Rosenzweig P-F Study administered shortly after a match as "possessing more superego" than the control group. Husman believes this increase in "superego" is to be expected. The boxers' aggressive behavior in athletic contests presumably made them feel guilty even though this aggression is socially sanctioned. After the boxing season their mean rating on Rosenzweig's "superego" trait tended to decrease (nonsignificantly) approaching that of a normal population. Husman does not report any significant differences for wrestlers.

Stone (**77**) also obtained indications of inhibited hostility following a socially sanctioned athletic contest. The TAT was given to football players both during and after the athletic season, and their responses

were compared with the responses of a matched control group. There was no difference between the two groups in fantasy aggression during the football season, but the football players showed significantly less manifest aggression on the TAT following the completion of the season. Interestingly enough, however, the aggression that the football players did display tended to be of a "projective" nature, i.e., it was attributed to an impersonal source. Stone argues that this aggression was projected from the self onto an impersonal source because of *S*'s aggression-anxiety. Clark (**12**), in a study reported earlier, also obtained a high proportion of "projective hostility" TAT responses and relatively few manifest aggression themes in an experimental condition presumably arousing guilt.

Thus, there is reason to believe that sanctioned overt aggression may lead to a reduction in the display of hostility, but not necessarily because of catharsis. The present studies suggest that aggression-anxiety frequently results (at least in middle class college students) from aggressive actions, even in situations where aggression is encouraged or permitted, and this anxiety may inhibit subsequent aggressive responses. Therefore, before an investigator can demonstrate catharsis unequivocally, it is necessary for him to show that any decrease in overt hostility is not due to the arousal of aggression-anxiety.

While the above studies have all the advantages inherent in dealing with "real-life" situations, they typically also suffer from inadequate controls. Ultimately, then, the crucial tests of the catharsis hypothesis must come from laboratory investigations utilizing experimentally aroused hostility. Several studies carried out under relatively controlled conditions have obtained results which, on the surface at least, are consistent with the notion of catharsis (**18, 22, 58, 74, 78**). However, there are plausible alternative explanations in all of these cases.

Thus, in the experiment by Pepitone and Reichling (**58**) mentioned earlier, the investigators angered all *S*s and then created conditions either facilitating or inhibiting the occurrence of overt aggression, i.e., high or low liking among the group members. The 13 two-man high liking groups at first expressed relatively strong hostility toward the instigator and then showed a steady decline in open aggression toward him. The 13 low liking groups, in contrast, did not show this decline. Furthermore, in accord with the catharsis hypothesis, the high liking groups, which had expressed a reliably greater total volume of aggression toward the instigator, rated him significantly more favorably at the conclusion of the experiment than the low liking groups. Here again it is possible that the relatively strong hostility expressed by the members of the high liking groups had evoked strong guilt feelings (or aggression-anxiety). These feelings could have (*a*) inhibited the later aggressive responses and (*b*) produced the favorable ratings of the instigator. In addition, as Pepitone and Reichling suggest on the basis of (already cited) findings by Thibaut and Coules (**79**), the restraints against aggressive behavior felt by the members of the low liking groups could have added to the hostility-producing frustration. Guilt or aggression-anxiety can also account for French's observation that the groups that had previously expressed the most aggression tended to be the least aggressive in the next situation (**22**, p. 288).

Somewhat similar explanations, plus others, may be offered for Thibaut's results (**78**). After frustrating the status aspirations of 18 five- or six-person groups, he succeeded in having these *S*s direct some aggression against the high status teams. Nine of the low status teams then won high status while the other nine remained low in status. Since the frustration does not persist in the former groups, they are the only ones who should demonstrate a cathartic reduction of hostility. For these teams the correlation between pre- and postupward mobility overt aggression is $-.69(p=.05)$. For the persistently frustrated teams remaining in the low status position this r is only $-.22$.

This difference in correlations can be interpreted either according to the catharsis formulation or in terms of a quite different process discussed more recently by Thibaut and his colleagues (**80**, pp. 95–97). They have shown that the instigation to aggression is reduced when an individual's hostile action appears to further his ends. Thus, the decrease in aggressive behavior observed in the successfully upward-mobile low status teams could have resulted from the *S*'s judgment that this behavior had been at least partly responsible for their rise in status. It also is possible that the group members who previously had expressed a good deal of hostility had developed guilt feelings about this when their frustration was shown to be only temporary (i.e., when they were elevated to the high status position). These guilt feelings then could have produced the decline in overt hostility as well as the second catharsis-like finding noted by Thibaut; the successfully upward-mobile low status teams did not give less favorable sociometric ratings to the other, high status teams in the second part of the experiments (after their status rise) as did the persistently low status groups. The former groups could have rated the teams they displaced relatively more favorably because of guilt created by their previous aggression.

In one of the most recent investigations of catharsis, Feshbach (**18**) tested the hypothesis that the instigation to aggression could be reduced through symbolic satisfactions. Two experimental treatments were applied to introductory psychology classes angered by an insulting lecturer: (*a*) interpolation of fantasy activity (administration of four TAT cards) before having the *S*s respond to the two main aggression measures—a brief questionnaire assessing attitudes toward the experiment, and the Sentence Completion Test; and (*b*) interpolation of nonfantasy activity before the *S*s responded to these forms. A control group of *S*s was given the interpolated fantasy task without the insult treatment. Both aggression instruments proved to be sensitive to aggression arousal, with the insulted fantasy group scoring significantly lower on these measures than the insulted nonfantasy group. The insulted fantasy group also expressed significantly more total aggression to the TAT cards than the noninsulted fantasy group. Further support for the catharsis hypothesis is found in the significant negative r between TAT aggression and aggression on the questionnaire ($r=-.25$, $p=.01$) in the insulted group. Feshbach presents some inconclusive evidence suggesting that the lowered aggression in the insulted fantasy group was not due to guilt arousal (his guilt measures had not been validated), and that the greater aggression in

the control group was not due to the frustrating qualities of the interpolated neutral task.

However, the previously mentioned findings reported by Thibaut and Coules (**79**) cast some doubt on Feshbach's last contention. These writers have shown that the type of interference with the occurrence of aggressive responses brought about by the administration of the neutral task could have increased the aggressive drive. McClelland, maintaining his previously expressed (**43**) views against the notion of drive satisfaction through implicit symbolic responses, also questions Feshbach's interpretations (**45**, p. 53). He suggests that the instructions for the TAT ("feel free to write whatever you like . . . ") reduced *Ss*' irritation with the insulting *E*. Along these lines, the present writer has observed that frustrated *Ss* may find the comparatively novel TAT test sufficiently enjoyable to overcome much of their frustration-produced annoyance.[7]

Evaluation

The catharsis hypothesis is not accepted by all personality theorists (e.g., **1**, **44**). Allport (**1**), for example, disputes the concept, implicit in many discussions of catharsis, of a "reservoir" of hostility that may be "drained off" in any number of different ways. In opposition to this "drainage theory," he cites studies (e.g., **75**) reporting positive correlations among different modes of aggression instead of the negative correlations explicity predicted by Dollard, Doob, et al. (**17**, p. 51). In general, Allport suggest, "All this evidence is hard on the theory that free-floating aggression may be 'drained off' from one object to another. . . . It simply is not true that a given quantum of free-floating aggression can be used up in this, that, or the other way" (**1**, p. 359).

Dollard and his colleagues were aware of findings similar to that reported by Allport, (but discuss them only in a footnote to p. 52):

> It appears that there are positive correlations between the occurrence of various forms of overt aggression and between various forms of non-overt aggression. The reciprocal relationship is probably between overt and non-overt on the one hand and between self-directed and object-directed on the other and may not appear at all if the amount of frustration is not held constant.

There is evidence both for and against aspects of this hypothesized negative relationship between different modes of aggression. McClelland cast some doubt on the expectation of an inverse relation between overt and covert hostility (**44**, pp. 516-517). He notes that in Stone's study of sanctioned overt aggression (**77**) there were many cases of football players who exhibited high aggression both overtly on the field and, covertly, in response to the TAT, or who were low in both. The previously cited results obtained by Lesser (**35**) and Mussen and Naylor (**54**) are in accord with this observation. However, there also are recent experimental findings (**80**, p. 113) showing that angered *Ss* who attack others tend reliably not to demonstrate self-aggression, supporting the "slender thread of evidence" reported by the Yale psychologist in defense of their hypothesis (**17**, p. 52).

Again, the above evidence is inconclusive with regard either to the

[7] Feshbach did not have equal sex ratios in all conditions. Preliminary evidence collected by the present writer suggests that men "get over" their anger much more quickly than women after working on interpolated tasks.

confirmation or rejection of the catharsis hypothesis. The reported positive correlations among various modes of aggression could well stem from generalized habits. Studies of social prejudice have shown repeatedly that hostility is a broadly generalized response for the prejudiced individual. Strangers, minority group members, foreigners and all who are "different," tend to be equated as potentially dangerous stimulus objects and all arouse aggressive responses within him. Thus, the person who is consistently aggressive, overtly and covertly, might well have learned to tie many objectively different kinds of situations together on a single stimulus generalization gradient. All of these situations for example, might be labeled by him as threatening, with an instigation to aggression resulting from the labeling response.

In general, then, a major difficulty with these studies, regardless of whether they show positive or negative relationships among various modes of aggression, is that they are correlational in nature. To use the negative relationship cited above as an illustration, we cannot prove that the self-aggression subsequently reduced the instigation to other-aggression and/or vice versa. The negative correlation may be due to some common process acting upon both types of responses. Self-aggression and other-aggression may be relatively stable and incompatible response patterns through past learning so that any given individual is likely either to attack himself or another when frustrated. This suggestion, of course, is the basis of Rosenzweig's frustration-response typology (**61**).

Experimental investigations can overcome this type of difficulty if designed properly, i.e., if there is cognizance of the conditions limiting the generality of the catharsis hypothesis. As was pointed out earlier, an adequate test of this hypothesis must: (*a*) determine whether there is inhibited aggression and, if possible, eliminate aggression-anxiety, (*b*) protect the *S*s from continuing frustration, (*c*) demonstrate that a greater volume of aggressive behavior was expressed in the "cathartic" condition than in the control condition, and (*d*) analyze separately the results for the *S*s who have a "habit" of behaving aggressively when faced by frustration (and, consequently, who would persist in this type of response long after other *S*s have ceased acting in a hostile manner).

The term, "catharsis," also has been applied to a phenomenon somewhat different from that discussed by the F-A psychologists. McClelland and Apicella (**42**), for example, differentiate between anger and aggression, and they propose that the emotion, anger, has "some tension-reducing capacity of its own *prior* to the discovery of some object against which it may be directed" (cited in **44**, p. 513). Sears, Maccoby, and Levin refer to a similar process when they suggest that "adults often experience this relief of tension after they have lost their tempers—they speak of 'letting off steam'. . . ." (**71**, p. 225). However, they believe the emotional tension results from the approach-avoidance conflict of wanting to injure but being afraid or reluctant to do so. Whatever the source of this tension, this usage (reduction of discomfort) is not identical with the hypothesis formulated by Dollard, Doob, et al. (which involves a decrease in the aggressive drive).

Recent experimental evidence pro-

vided by Worchel (**83**) indicates that the expression of aggression can reduce performance-interfering tension. After aggressing against the frustrating experimenter either directly to him or indirectly to his assistant, college students performed significantly better on a digit-symbol test than the insulted *Ss* not given this opportunity for "catharsis." Worchel assumes that the performance of the latter *Ss* was disrupted by their anxiety and aggression, while these interfering emotional states were dispelled in the cathartic treatment. Along these lines, we might speculate whether the blocking of aggressive responses in the noncatharsis conditions was one of the frustrations interfering with effective performance. If so, the performance difference might be due to this latter type of frustration predominantly. At any rate, the reduction of performance-disruptive emotional states following the expression of hostility does not necessarily mean that the instigation to aggression also was reduced.

REFERENCES

1. ALLPORT, G. W. *The nature of prejudice.* Cambridge, Mass.: Addison-Wesley, 1954.
2. ANASTASI, ANNE, COHEN, NADIS, & SPATZ, DOROTHY. A study of fear and anger in college students through the controlled diary method. *J. genet. psychol.*, 1948, **73**, 243–249.
3. APPEL, M. H. Aggressive behavior of nursery school children and adult procedures in dealing with such children. *J. exp. educ.*, 1942, **2**, 195–199.
4. BACH, G. R. Young children's play fantasies. *Psychol. Monogr.*, 1945, **59**, No. 2 (Whole No. 272).
5. BARKER, R. G., DEMBO, TAMARA, & LEWIN, K. Frustration and regression: an experiment with young children. *Univer. Iowa Stud. Child Welf.*, 1941, **18**, 1–314.
6. BATESON, G. The frustration-aggression hypothesis and culture. *Psychol. Rev.*, 1941, **48**, 350–355.
7. BLOCK, JEANNE, & MARTIN, B. C. Predicting the behavior of children under frustration. *J. abnorm. soc. psychol.*, 1955, **51**, 281–285.
8. BORNSTON, FRIEDA L., & COLEMAN, J. C. The relationship between certain parents' attitudes toward child rearing and the direction of aggression of their young adult offspring. *J. clin. psychol.*, 1956, **12**, 41–44.
9. BROWN, J. S., & FARBER, I. E. Emotions conceptualized as intervening variables —with suggestions toward a theory of frustration. *Psychol. Bull.*, 1951, **48**, 465–495.
10. CHASDI, ELEANOR H., & LAWRENCE, MARGARET S. Some antecedents of aggression and effects of frustration in doll play. In D. McClelland, (Ed.) *Studies in motivation.* New York: Appleton-Century-Crofts, 1955.
11. CHILD, I. L. Socialization. In G. Lindzey (Ed.), *Handbook of social psychology*, Vol. 2, Ch. 18. Cambridge, Mass: Addison-Wesley, 1954.
12. CLARK, R. A. The effect of sexual motivation on phantasy. In D. McClelland (Ed.), *Studies in motivation.* New York: Appleton-Century-Crofts, 1955.
13. COHEN, A. R. Social norms, arbitrariness of frustration, and status of the agent of frustration in the frustration-aggression hypothesis. *J. abnorm. soc. Psychol.*, 1955, **51**, 222–226.
14. DAVITZ, J. R. The effects of previous training on postfrustration behavior. *J. abnorm. soc. Psychol.*, 1954, **47**, 309–315.
15. DINWIDDIE, F. M. *An application of the principle of response generalization to the prediction of displacement of aggressive responses.* Washington: The Catholic Univer. of America Press, 1955.
16. DOLLARD, J. Hostility and fear in social life. *Social forces*, 1938, **17**, 15–25.
17. DOLLARD, J., DOOB, L. W., MILLER, N. E. MOWRER, O. H., & SEARS, R. R. *Frustration and aggression.* New Haven: Yale Univer. Press, 1939.
18. FESHBACH, S. The drive-reducing function of fantasy behavior. *J. abnorm. soc. Psychol.*, 1955, **50**, 3–11.
19. FESHBACH, S. The catharsis hypothesis and some consequences of interaction with aggressive and neutral play objects. *J. Personal.* 1956, **24**, 449–462.
20. FESHBACH, S., & SINGER, R. The effects

of personal and shared threats upon social prejudice. *J. abnorm. soc. Psychol.*, 1957, **54**, 411–416.

21. FREDERIKSEN, N. The effects of frustration on negativistic behavior of young children. *J. genet. Psychol.*, 1942, **61**, 203–226.
22. FRENCH, J. R. P. Organized and unorganized groups under fear and frustration. In K. Lewin, C. E. Meyers, J. Kalhorn, & M. L. Farber (Ed.), *Authority and frustration. Univer. Iowa Studies in Child Welfare*, Iowa City: Univer. of Iowa Press, 1944.
23. GATLING, F. P. Frustration reactions of delinquents using Rosenzweig's classification system. *J. abnorm. soc. Psychol.* 1950, **45**, 749–752.
24. GRAHAM, F. K., CHARWAT, W. A., HONIG, A. S., & WELTZ, P. C. Aggression as a function of the attack and the attacker. *J. abnorm. soc. Psychol.*, 1951, **46**, 512–520.
25. HANER, C. F., & BROWN, P. A. Clarification of the instigation to action concept in the frustration-aggression hypothesis. *J. abnorm. soc. Psychol.*, 1955, **51**, 204–206.
26. HIMMELWEIT, HILDE. Frustration and aggression: A review of recent experimental work. In T. H. Pear, (Ed.), *Psychological factors of peace and war*, New York: Philosophical Library, 1950.
27. HOKANSON, J. E., & GORDON, J. E. The expression and inhibition of hostility in imaginative and overt behavior. *J. abnorm. soc. Psychol.*, in press
28. HOLZBERG, J. D., BURSTEN, B., & SANTICCIOLI, A. The reporting of aggression as an indication of aggressive tension. *J. abnorm. soc. Psychol.*, 1955, **50**, 12–18.
29. HUSMAN, B. F. Aggression in boxers and wrestlers as measured by projective techniques. *Res. Quart. Amer. Ass. Hlth. Phys. Educ.*, 1955, **26**, 421–425.
30. JOHNSON, W. R., & HUTTON, D. C. Effects of a combative sport upon personality dynamics as measured by a projective test. *Res. Quart. Amer. Ass. Hlth. Phys. Educ.*, 1955, **26**, 49–53.
31. KAGAN, J. The measurement of overt aggression from fantasy. *J. abnorm. soc. Psychol.*, 1956, **52**, 390–393.
32. KEISTER, M. E., & UPDEGRAFF, R. The behavior of young children in failure: an experimental attempt to discover and to modify undesirable responses of pre-school children to failure. *Univer. Iowa Stud. Child Welfare*, 1938, **14**, 27–82.
33. KELLEY, H. H. Communication in experimentally created hierarchies. *Hum. Relat.*, 1951, **4**, 39–56.
34. KENNY, D. T. *An experimental test of the catharsis theory of aggression.* Ann Arbor: Univer. Microfilms, 1953.
35. LESSER, G. S. The relationship between overt and fantasy aggression as a function of maternal response to aggression. *J. abnorm. soc. Psychol.*, 1957, **55**, 218–221.
36. LEVIN, H., & SEARS, R. R. Identification with parents as a determinant of doll play aggression. *Child Develpm.*, 1956, **27**, 135–153.
37. LEVIN, H., & TURGEON, VALERIE F. The influence of mother's presence on children's doll play aggression. *J. abnorm. soc. Psychol.* 1957, **55**, 304–308.
38. LINDZEY, G. An experimental examination of the scapegoat theory of prejudice. *J. abnorm. soc. Psychol.*, 1950, **45**, 296–309.
39. LIVSON, N., & MUSSEN, P. H. The relation of ego control to overt aggression and dependency. *J. abnorm. soc. Psychol.*, 1957, **55**, 66–71.
40. MARX, M. H. Some relations between frustration and drive. In M. R. Jones, (Ed.), *Nebraska symposium on motivation*, **1956**, Lincoln, Neb.: Univer. Nebraska Press, 1956.
41. MASLOW, A. H. Deprivation, threat and frustration. *Psychol. Rev.*, 1941, **48**, 364–366.
42. MCCLELLAND, D. C., & APICELLA, F. S. A functional classification of verbal reactions to experimentally induced failure. *J. abnorm. soc. Psychol.*, 1945, **40**, 376–390.
43. MCCLELLAND, D. C., CLARK, R. A., ROBY, T. B., & ATKINSON, J. W. The projective expression of needs: IV. The effect of need for achievement on thematic apperception. *J. exp. Psychol.*, 1949, **39**, 242–255.
44. MCCLELLAND, D. C. *Personality.* New York: Dryden, 1951.
45. MCCLELLAND, D. C. Personality. In P. R. Farnsworth, and Q. McNemar, (Eds.) *Ann. Rev. Psychol.*, Vol. 7, Stanford: Annual Reviews, 1956.
46. MCKEE, J. P., & LEADER, FLORENCE, B. The relationship of socioeconomic status and aggression to the competitive behavior of pre-school children. *Child Develpm.*, 1955, **26**, 135–142.

47. McKellar, P. The emotion of anger in the expression of human aggressiveness. *Brit. J. Psychol.*, 1949, **39**, 148–155.
48. Miller, D. R., & Swanson, G. E. The study of conflict. In M. R. Jones (Ed.), *Nebraska symposium on motivation, 1956*, Lincoln, Neb.: Univer. Nebraska Press, 1956.
49. Miller, N. E. The frustration-aggression hypothesis. *Psychol. Rev.*, 1941, **48**, 337–342.
50. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.
51. Mintz, A. A re-examination of correlations between lynchings and economic indices. *J. abnorm. soc. Psychol.*, 1946, **41**, 154–160.
52. Morlan, G. K. A note on the frustration-aggression theories of Dollard and his associates. *Psychol. Rev.*, 1949, **56**, 1–8.
53. Murney, R. G. *An application of the principle of stimulus generalization to the prediction of object displacement*, Washington, D. C.: Catholic Univer. of America Press, 1955.
54. Mussen, P. H., & Naylor, H. K. The relationships between overt and fantasy aggression. *J. abnorm. soc. Psychol.*, 1954, **49**, 235–240.
55. Newcomb, T. M. Autistic hostility and social reality. *Hum. Relat.* 1947, **1**, 3–20.
56. Otis, N. B., & McCandless, B. Responses to repeated frustrations of young children differentiated according to need area. *J. abnorm. soc. Psychol.*, 1955, **50**, 349–353.
57. Pastore, N. The role of arbitrariness in the frustration-aggression hypothesis. *J. abnorm. soc. Psychol.*, 1952, **47**, 738–731.
58. Pepitone, A., & Reichling, G. Group cohesiveness and the expression of hostility. *Hum. Relat.*, 1955, **8**, 327–337.
59. Pepitone, A., & Kleiner, R. The effect of threat and frustration on group cohesiveness. *J. abnorm. soc. Psychol.* 1957, **54**, 192–199.
60. Rosenzweig, S. The experimental measurement of types of reactions to frustration. In Murray, H. A. *Explorations in personality*, New York: Oxford Press, 1939.
61. Rosenzweig, S. An outline of frustration theory. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald, 1944.
62. Rosenzweig, S., & Rosenzweig, L. Aggression in problem children and normals as evaluated by the Rosenzweig P-F study. *J. abnorm. soc. Psychol.*, 1952, **47**, 683–687.
63. Rotter, J. B., & Wickens, D. D. The consistency and generality of ratings of "social aggressiveness" made from observation of role playing situations. *Amer. Psychologist*, 1947, **2**, 333.
64. Sargent, S. S. Reaction to frustration—a critique and hypothesis. *Psychol. Rev.*, 1948, **55**, 108–114.
65. Sears, Pauline S. Doll play aggression in normal young children: Influence of sex, age, sibling status, father's absence. *Psychol. Monogr.*, 1951, **65**, No. 6 (Whole No. 323).
66. Sears, R. R. Non-aggressive reactions to frustration. *Psychol. Rev.*, 1941, **48**, 343–346.
67. Sears, R. R. Effects of frustration and anxiety on fantasy aggression. *Amer. J. Orthopsychiatry*, 1951, **21**, 498–505.
68. Sears, R. R. A theoretical framework for personality and social behavior. *Amer. Psychologist*, 1951, **6**, 476–483.
69. Sears, R. R., Pintler, Margaret H., & Sears, Pauline S. Effect of father separation on preschool children's doll play aggression. *Child. Develpm.*, 1946, **17**, 219–243.
70. Sears, R. R., Whiting, J. W. M., Nowlis, V., & Sears, Pauline S. Some child-rearing antecedents of aggression and dependency in young children. *Genet. Psych. Monogr.*, 1953, **47**, 135–234.
71. Sears, R. R., Maccoby, Eleanor E., & Levin, H. *Patterns of child rearing*, Evanston, Ill.: Row, Peterson, 1957.
72. Seward, J. P. Aggressive behavior in the rat: I. General characteristics; age and sex differences. *J. comp. Psychol.*, 1945, **38**, 175–197.
73. Seward, J. P. Aggressive behavior in the rat: III. The role of frustration. *J. comp. Psychol.*, 1945, **38**, 225–238.
74. Seward, J. P. Aggressive behavior in the rat: IV. Submission as determined by conditioning, extinction and disuse. *J. comp. Psychol.*, 1946, **39**, 51–76.
75. Stagner, R. Studies of aggressive social attitudes: I. Measurement and interrelation of selected attitudes. *J. soc. Psychol.*, 1944, **20**, 109–120.
76. Stagner, R., & Congdon, C. S. Another failure to demonstrate displacement of

aggression. *J. abnorm. soc. Psychol.*, 1955, **51**, 695–696.

77. STONE, A. A. *The effect of sanctioned overt aggression on total instigation to aggressive responses.* Unpublished honors thesis, Harvard Univer., 1950.
78. THIBAUT, J. An experimental study of the cohesiveness of underprivileged groups. *Hum. Relat.*, 1950, **3**, 251–278.
79. THIBAUT, J., & COULES, J. The role of communication in the reduction of interpersonal hostility. *J. abnorm. soc. Psychol.*, 1952, **47**, 770–777.
80. THIBAUT, J., & RIECKEN, H. Authoritarianism, status, and the communication of aggression. *Hum. Relat.*, 1955, **8**, 95–120.
81. WRIGHT, G. O. Projection and displacement: a cross-cultural study of folk-tale aggression. *J. abnorm. soc. Psychol.*, 1954, **49**, 523–528.
82. WRIGHT, M. E. The influence of frustration upon the social relations of young children. *Character and Pers.*, 1943, **12**, 111–122.
83. WORCHEL, P. Catharsis and the relief of hostility. *J. abnorm. soc. Psychol.*, 1957, **55**, 238–243.
84. YARROW, L. J. The effect of antecedent frustration on projective play. *Psychol. Monogr.*, 1948, **62**, No. 6 (Whole No. 293).
85. ZANDER, A. F. A study of experimental frustration. *Psychol. Monogr.*, 1944, **56**, No. 3 (Whole No. 256).

Received January 29, 1958.

PSYCHOLOGICAL BULLETIN
VOL. 55, NO. 5, 1958

SECONDARY REINFORCEMENT: A REVIEW OF RECENT EXPERIMENTATION[1]

JEROME L. MYERS
University of Massachusetts

In the past decade, psychology has witnessed two important research trends in the area of secondary reinforcement. Whereas the bulk of previous research[2] was concerned with demonstrating the existence of secondary reinforcement, recent experimentation has centered about (*a*) the question of defining secondary reinforcement and, in particular, exploring alternative explanations of the data and (*b*) the investigation of parameters which effect various measures of secondary reinforcement. Some experimenters have been concerned with differentiating between the role of the originally neutral stimulus as a cue and as a reinforcer (**9, 30, 33, 43**). Others have been interested in testing the applicability of a discrimination hypothesis (**3, 25**). Much experimentation has dealt with the effects of such variables as drive (**7, 13, 14, 20, 27, 36, 38, 40, 41, 42**), frequency of pairing of the neutral and primary reinforcing stimuli (**2, 19, 27**), reinforcement schedules (**5, 10, 11, 24, 25, 29, 30**), and amount of reinforcement (**6, 21, 23**).

This paper is a review of recent experimentation dealing with the problems outlined above. An attempt has been made to integrate and evaluate the results of a number of experiments, to point out areas of agreement and disagreement among experments, and to suggest the directions future research should take.

[1] The author wishes to thank N. A. Myers for her helpful comments on the manuscript.

[2] See Miller's review of the literature (**28**).

A DISCRIMINATIVE STIMULUS HYPOTHESIS

A study of Schoenfeld, Antonitis, and Bersh (**37**) has been instrumental in redefining a secondary reinforcer and in focusing attention upon the possible equivalence of secondary reinforcers and discriminative stimuli. Two groups of rats were given food pellet reinforcements for bar pressing responses. For the experimental group a light of one-second duration went on at the onset of eating, rather than simultaneous with bar pressing and prior to eating as is usual in secondary reinforcement experiments. No light stimulus was presented to the control group. The measure of secondary reinforcement was the rate of bar pressing during extinction trials when the light alone was presented. The lack of a significant difference between the rates of the two groups led the authors to conclude that the neutral stimulus must precede the primary reinforcer if secondary reinforcement is to occur.

To explain their results, Schoenfeld, et al. put forth the hypothesis that the efficacy of the originally neutral stimulus during extinction trials is dependent upon its serving as a discriminative stimulus during training. A discriminative stimulus is one whose presence is a cue for a particular response, and whose absence is a cue for not responding. If, for example, bar pressing in the presence of light is rewarded while bar pressing in the absence of light is not rewarded, light becomes a discrimina-

tive stimulus. Since in the experiment under discussion (in which secondary reinforcement was not obtained) the conditions for the light becoming a discriminative stimulus did not hold, the authors concluded that the neutral stimulus must be established as a discriminative stimulus if it is to function as a secondary reinforcer.

Dinsmoor (**9**) designed an experiment to further investigate the relationship between secondary reinforcers and discriminative stimuli. He gave his animals extensive discrimination training in the Skinner box. They were then divided into three groups, each of which was extinguished under one of three conditions. One group was extinguished with the discriminative stimulus, a light, continuously present except during three-second intervals following each bar press. The difference between the extinction response rate of this group and the extinction response rate of a control group was a measure of the role of the discriminative stimulus. A second group had to press the bar to evoke the discriminative stimulus for three seconds. Here, the difference between the extinction response rate of this group and that of the control group was a measure of secondary reinforcement. The control group was extinguished in the complete absence of the discriminative stimulus. The cumulative response curves for the first two groups virtually coincided and were consistently higher than that of the control group. Dinsmoor pointed out that the secondary reinforcer is a stimulus whose presence is contingent upon the response, while the discriminative stimulus is presented without regard to the subject's pattern of response. He concluded that the distinction between the reinforcing and discriminative functions of a stimulus can only be made in terms of this difference in temporal schedules for presenting the stimulus.

An experiment by Ratner (**33**) appears to present a challenge to the discriminative stimulus hypothesis. A group of rats was trained to approach a water dipper at the sound of a click. After this training, water was removed and a bar was inserted into the apparatus and, for half of the animals, a click followed each bar press. Both goal approaching and bar pressing responses were counted. A significant difference in the number of presses for the click and no-click groups was obtained, but there was little difference in the number of goal approaching responses or in the number of bar presses followed by goal approaching responses. Ratner concluded that, despite the acquisition of secondary reinforcing properties, the click was not a discriminative stimulus for goal approaching during extinction trials.

Ratner's brief report of his study did not permit an extensive presentation of the training procedure, but his statement that the rats were trained to run to a dipper at the sound of a click strongly implies that at that point in the experiment the click was a discriminative stimulus for goal approaching. Therefore, we must conclude that the click lost its discriminative property while becoming a reinforcer for a new response, bar pressing. The results suggest that a stimulus may be reinforcing, although lacking discriminative properties. It cannot be concluded, however, that discrimination training is not a necessary condition for the establishment of secondary reinforcers. In view of its implications

for Schoenfeld's hypothesis (**37**) in particular, and for secondary reinforcement theories in general, the study certainly merits replication.

A recent experiment by Wyckoff, Sidowski, and Chambliss (**43**) casts doubt not only upon the discriminative stimulus hypothesis, but upon the general concept of secondary reinforcement as well. Wyckoff et al. point out that if the bar and the reinforcement magazine are close together, any stimulus which has become a discriminative stimulus for goal approaching will tend to keep the subject in the vicinity of the bar, thus increasing the probability that the bar will be pressed. In the apparatus used by Wyckoff et al. the bar was placed at the opposite end of the box from the reinforcement magazine. The rats were initially trained to approach and lick a dipper at the sound of a buzzer. During the subsequent test session, the buzzer was presented to the experimental group after each bar press, and to the control group after each 10-second interval of nonbar pressing. Despite the fact that the buzzer was firmly established as a discriminative stimulus, and despite the fact that the situation was designed so that one group would be reinforced for bar pressing and the other group for not bar pressing, no difference was obtained in rate of bar pressing.

These results are highly suggestive. It is possible that discrimination training is necessary to establish a secondary reinforcer for two reasons: (*a*) The discriminative stimulus keeps the animal in the vicinity of the bar, and (*b*) the discriminative stimulus keeps the animal active, again increasing the probability of the appropriate response. As Wyckoff points out, other studies involving other apparatuses and procedures have yielded secondary reinforcement effects. However, the data gathered in a large number of Skinner box studies are now open to serious doubt, and it is evident that these experiments should be repeated with the apparatus modifications employed by Wyckoff et al.

The paucity of experimental evidence prevents any meaningful evaluation of the discriminative stimulus hypothesis, at present. Certainly, it cannot be concluded that discrimination training is a necessary requirement for establishing a secondary reinforcer, particularly when the secondary reinforcer is defined in terms of learning a new response rather than in terms of resistance to extinction. If discrimination training is necessary, it is still not clear whether the originally neutral stimulus must retain its discriminative properties during the test session if it is to serve as a secondary reinforcer. Ratner's results suggest a negative answer to this, but for conclusive evidence more experiments utilizing measures of goal approaching behavior, such as number and latency of responses, are required. Finally, in light of Wyckoff's results, the possibility should be considered that the apparent role of discrimination training in establishing secondary reinforcers is, at least in part, a function of the apparatus design. This will require further experimental verification.

A Discrimination Hypothesis

Two studies by Bitterman and his associates (**3**, **12**) have led to the formulation of a discrimination hypothesis of secondary reinforcement. Both experiments involved the same basic design, the major difference being in the length of intertrial intervals

employed. Rats were required to run the length of a runway and then jump to a goal box. On rewarded trials the interior of the goal box was black and on nonrewarded trials it was white. On extinction runs, half of the animals ran to the black interior while the other half ran to the previously nonrewarded white interior. A secondary reinforcement approach must lead to a prediction of shorter running times for the first group. However, in both experiments, the group tested in the presence of the previously nonrewarded color was superior. Conditioning and extinction are similar for this group in that in both phases of the experiment they are consistently not rewarded when a particular color is present. On the basis of these results, it was suggested that a secondary reinforcement principle is not sufficient to explain resistance to extinction, and that resistance to extinction is a function of the similarity between conditions of training and extinction. A number of the experiments in the next section should facilitate an evaluation of the discrimination hypothesis and an understanding of its implications for research in this area.

Schedules of Reinforcement

Dinsmoor, Kish, and Keller (**11**) gave two groups of rats discrimination training in a Skinner box, then presented the discriminative stimulus, a light, to one group after each response in the dark, and to the second group periodically, that is, only after the first response in each five minute dark period. Rate of responding was higher in the regular reinforcement group in the early sessions, but the periodic reinforcement group showed a higher response rate by the fourth session. The authors concluded that regular and periodic reinforcement do not produce different total effects. The generality of this conclusion must remain in question. It is possible that the use of other procedures to establish the light as a secondary reinforcer, or the use of periods other than five minutes in duration might yield different results, presumably an advantage for the periodic technique. In any event, it would be of interest to investigate the effects of other schedules of secondary reinforcement.

Dinsmoor (**10**) has also explored the effects of periodic schedules of primary reinforcement upon the establishment of secondary reinforcers. One group of rats was given discrimination training and its response rate was then measured in alternate light and dark periods, with food withheld. A second group was given the same training but the prevalent test condition was darkness, each bar press being followed by three seconds of light. Food reward was presented periodically to both these groups in the presence of light during training. A third group was regularly reinforced for responses in the presence of light and was tested in the same manner as the first group. The periodic reinforcement groups showed greater response rates during the test session than did the regular reinforcement group. The secondary reinforcement group, which was "rewarded" by the three-second light, showed a cumulative response curve similar to that of the periodic reinforcement group tested in alternate light and dark phases. The results thus substantiate the hypothesis of a functional relation between secondary reinforcers and discriminative stimuli and suggest the importance of schedules of primary

reinforcers in the establishment of secondary reinforcers.

Clayton (**5**) has recently investigated the effects of regular and intermittent reinforcement, both primary and secondary, upon the strength of secondary reinforcement. Her rats were reinforced for bar pressing in the presence of light according to one of three schedules: 150 regular reinforcements, 150 reinforcements given over 250 trials (a trial being defined as the presentation of a light which goes off when the response is made), and 100 reinforcements given over 150 trials. Each of these groups was then extinguished in one of two ways: presentation of the light after each bar press or presentation of the light after 60% of the bar presses. No significant effects were obtained. Comparison of the six experimental groups with a control group which was extinguished in the absence of light showed significantly more responses for the experimental groups. The absence of effects due to reinforcement schedule is somewhat disturbing since it is reasonable to expect that (*a*) intermittent primary reinforcement should increase resistance to extinction and (*b*) if the light is a reinforcing stimulus, its intermittent presentation during extinction sessions should increase resistance to extinction. Nor are Clayton's results on primary reinforcement scheduling in accordance with expectations based on the Dinsmoor (**10**) study of the effects of primary periodic reinforcement. It is possible that two days of discrimination training are not sufficient to establish differences between continuous and intermittent reinforcement groups, and that the rats extinguished too rapidly for the intermittent secondary reinforcement to have any appreciable effect upon response rate.

Saltzman (**35**) trained rats on a runway, rewarding half the rats in a white goal box, the other half in a black goal box. He then tested the rats in a **U** maze, forcing them to choose between a white and black goal box, with no food reward present in either box. Groups given alternate rewarded and nonrewarded trials in the runway learned the maze problem as well as groups working for food, and made significantly fewer errors than secondary reinforcement groups given only rewarded runway trials. It appears that partial reinforcement during training is more effective than continuous reinforcement in establishing secondary reinforcers. This conclusion is supported by Notterman's (**30**) finding that secondary reinforcement is a monotonically increasing function of the number of unrewarded trials given during training.

McClelland and McGown (**24**) designed an experiment to approximate the variable training which occurs in the nonlaboratory situation, in an attempt to establish less transient secondary reinforcement effects than generally have been shown. One group of rats (the specific-reinforcement group) was required to enter a circular goal alley, and turn left. Food was always found in the same place, a fixed distance to the left of the entry, and in front of a barrier (thus preventing a rat from approaching the food from the right). A second group (the general-reinforcement group) was rewarded in four different parts of the alley in a random order. Since no barrier was present they were permitted both right and left turns from the entry. Furthermore, on half the trials these rats found the pellets on the floor of

the alley, while on the other half of the trials, the rats were required to stop in the "correct" section, before a food pellet was awarded to them. To test the secondary reinforcement value of the goal alley, a **T** maze was placed at the entrance, and the rats were placed at one end of the horizontal bar and given the choice of running straight ahead or turning into the vertical bar and into the alley. During this test, the general-reinforcement group showed quicker entry into the goal alley, and less learning errors than the specific-reinforcement group. At the end of 25 trials, the general-reinforcement group showed no decrement in speed of entering the alley. These results suggest that increasing variability in the conditions present during training hinders the extinction of a secondary reinforcer.

Melching (**25**) varied the percentage of neutral stimulus presentations during Skinner box conditioning sessions. His rats were trained in one of three ways; either a buzzer always sounded following a bar press, or the buzzer sounded following 50% of the bar presses, or the buzzer never sounded. All bar presses were followed by food pellet delivery. Each of these groups was then extinguished in one of two ways; either with the buzzer following each bar press or without the buzzer presented at all. Melching found an overall significant difference in number of extinction bar presses between the groups. Since he unfortunately does not report the training or extinction conditions main effects or the interaction effect of the two, this overall F is difficult to interpret. However, t tests between all possible pairs of conditions yield one interesting comparison. No significant difference was found between the two groups conditioned with the buzzer following 50% of the bar presses. Secondary reinforcement theory would lead to a prediction of high extinction response rate for the group extinguished with the buzzer present. Melching accounts for these results in terms of a stimulus generalization hypothesis. For example, he suggests that stimulus generalization from 50% buzzer-during-conditioning to 100% buzzer-during-extinction was equivalent to stimulus generalization from 50% buzzer-during-conditioning to 0% buzzer-during-extinction. This would account for the lack of a difference in the extinction rates of the two groups. The results of other comparisons between groups are accounted for in a similar manner. Stimulus generalization hypothesis appears to be another label for Bitterman's discrimination hypothesis (**3**, **12**).

Myers (**29**) varied the percentage of candy reward (100%, 50%), the percentage of token reward during training (100%, 50%), and the percentage of token reward during extinction (100%, 0%). The subjects were 3–6 year-old children, and the measure of extinction was the rate of button pressing during extinction. Partial reinforcement on candy and on token led to more extinction responses than continuous reinforcement, and, contrary to Melching's (**25**) results, all four groups receiving the token during extinction made significantly more responses than those not receiving the token during extinction.

These investigations of reinforcement procedures merit attention for two reasons. Until it is experimentally demonstrated that conditions can be found under which secondary reinforcement appears to be effective

for more than 45 minutes in a Skinner box or 15 trials in a **U** maze, there is reason to doubt that much of human learning is motivated by secondary reinforcers. The Saltzman (**35**) and McClelland (**24**) studies are thus highly significant as indicators of ways in which more potent secondary reinforcers can be established. Myers' *S*s, particularly the group receiving 50% token and 50% candy reinforcement, and the token during extinction, also showed little decline in response rate by the end of the extinction sessions. There is a need for further investigation of the effects of the percentage of primary and secondary reinforcement during training, the percentage of secondary reinforcement during extinction, and the schedule of presentation of both primary and secondary reinforcers.

The experiments reviewed in this section are also of interest because of their relationship to Bitterman's discrimination hypothesis. The Saltzman (**35**) and Myers (**29**) studies indicate that secondary reinforcement is something more than a result of the discriminability of test and training sessions. A discrimination theory would predict that a group running to a white goal box on both rewarded and nonrewarded trials during training would show greater learning in the **U** maze than a group running to a white goal box on rewarded trials and a black goal box on nonrewarded trials. The former group undergoes nonrewarded "white" trials in both training and test sessions, and should thus be less able to discriminate. Saltzman, however, found a difference in the opposite direction. A discrimination hypothesis leads to the prediction that Myers (**29**) should obtain results similar to those obtained by Melching (**25**). In Melching's terms, all those groups undergoing changes of equal magnitude in stimulus complex from conditioning to extinction should respond at about the same rate during extinction. However, all of Myers' groups receiving the token during extinction gave significantly more responses than those not receiving the token.

If certain experiments do not appear to support the discrimination hypothesis, there are still many results which are most simply described in discrimination terms. The Bitterman studies (**3**, **11**) fall in this category. One of the first secondary reinforcement studies, that of Bugelski (**4**), is also simply explained by a discrimination hypothesis. Two groups of rats receive a click and food following each bar press. One group is extinguished with the click, the other without. There is no need to postulate a secondary reinforcer; extinction is more similar to conditioning for the "click" group.

McClelland and McGown (**24**) have related their results to a discrimination viewpoint. They argue that the stimulus complex for the general-reinforcement group is proportionately less changed by the absence of food during the test than is the stimulus complex for the specific reinforcement group. This greater similarity (for the general-reinforcement group) of training and test sessions would then account for the superior performance of the general-reinforcement group.

Myers (**29**), on the basis of differences between token-during-extinction and no-token-during-extinction groups, has suggested a modification of the discrimination hypothesis. She is able to explain her significant main effects, and in fact the observed rank ordering of her eight experi-

mental groups, by assuming (*a*) that the token, through pairing with candy, has acquired secondary reinforcement properties and (*b*) that differences between per cent reinforcement in extinction and in conditioning are also of importance. Myers' view can be stated in the equation:

$$X = AY + BZ + C \qquad [1]$$

while the Bitterman-Melching approach may be expressed as

$$X = A|Y| + BZ + C \qquad [2]$$

where

X = total number of extinction responses
Y = the difference between percentage of secondary reinforcement in conditioning and extinction
Z = the difference between percentage of primary reinforcement in conditioning and extinction
$|Y|$ is an absolute value
A and B are slope constants

and C is the slope intercept constant, presumably to be identified with free operant level. Both Myers and Melching are concerned with differences between the stimulus complex in conditioning and extinction. The former also assumes a secondary reinforcement factor which leads her to take the direction of differences into account.

Research dealing with the ability of subjects to discriminate changes in the stimulus complex would seem to be a prerequisite to a reliable evaluation of the discrimination hypothesis. For example, it is quite possible that changes from 50% (during conditioning) to 100% (during extinction) buzzer or token presentation are not as readily discriminable as changes from 50% to 0%. Before we can adequately make quantitative predictions about extinction performance on the basis of either Equation [1] or [2], it may be necessary to scale subjective differences in percentage presentation, and to then restate our equations in terms of the subjective differences between conditioning and extinction.

At the present time it appears that some form of a discrimination hypothesis, possibly one which leaves room for the concept of secondary reinforcement, is promising. The approach has broad generality, being applicable to both Skinner box (**4**, **25**) and runway (**24**) experiments, as well as to studies in which resistance to extinction (**3**, **4**, **11**, **25**) is measured and studies in which the learning of a new response (**24**) is measured. The hypothesis takes into account the experimental conditions during both training and test sessions and facilitates predictions about learning or resistance to extinction in light of these conditions.

Frequency of Primary Reinforcement

It is reasonable to hypothesize that the effectiveness of secondary reinforcement is a function of the frequency of pairing of the primary reinforcer and the neutral stimulus during training. This hypothesis has been investigated in a number of studies. Bersh (**2**), using a light as the neutral stimulus, and bar press rate during extinction as the measure of secondary reinforcement, varied the number of paired presentations of food and light over a range of six points extending from 0 to 120. The only significant differences in extinction rates were between the 0 and 120 reinforcement groups and between

the 20 and 120 reinforcement groups. A plot of the median number of responses as a function of frequency of primary reinforcements yielded a negatively accelerated curve which appeared to have reached its asymptote, suggesting that increased frequency of reinforcement would have had little further effect. The results of this experiment might discourage the hypothesis that frequency of reinforcement is an important parameter in the secondary reinforcement field. However, the lack of significant differences between the controls (no pairings of primary reinforcer and neutral stimulus) and all groups except the most extreme (120 pairings) would suggest that Bersh's procedure was not an optimal one for establishing and measuring secondary reinforcement.

Corroboration of Bersh's results and additional information as well have come from an experiment by Miles (**27**). He also used a Skinner box and measured bar press rate during extinction, but he introduced a number of methodological innovations. The range of frequency of reinforcement covered six points from 0 to 160 reinforcements. A combination of light and click was paired with each food pellet presented. Extinction was carried on to a criterion rather than for a definite time interval. Perhaps the most important aspect of the design was the use of six nonsecondary reinforcement groups, conditioned in exactly the same way as the secondary reinforcement groups but extinguished without the light-click combination present. Plotting number of extinction responses as a function of frequency of reinforcement led to the finding that the secondary reinforcement groups showed greater resistance to extinction than the nonsecondary reinforcement groups at all data points, and that both curves were negatively accelerated and appeared to approach somewhat different asymptotes at about the same rate. Miles computed the ratio of nonsecondary reinforcement rate of responding to secondary reinforcement rate for each of the six data points. A plot of the six ratios against frequency of reinforcement revealed no marked deviation from a straight line. Miles therefore concluded that the presence of secondary reinforcement leads to the same proportionate increase in extinction response rate throughout the parameter.

Hall (**19**) tested **T** maze learning (white against black goal boxes) after the subjects had either 25, 50, or 75 reinforced runs on a straightaway to a goal box which was always either black or white for any one rat. The number of correct runs increased linearly with increased frequency of primary reinforcement, the 25 and 75 reinforcement groups being significantly different. Extrapolation to zero reinforcements suggests that the response-frequency of reinforcement function is negatively accelerated as both Bersh (**2**) and Miles (**27**) found.

Further research is required to more precisely determine the shape of the performance-frequency of reinforcement function. Well-controlled and extensive experiments, utilizing the learning of a new response as a measure of secondary reinforcement, would also be welcome. Considering the lack of adequate controls and the limited range of values explored, the Hall (**19**) experiment can only be considered a first step in that direction.

Temporal Intervals

Bersh (**2**) varied the duration of a light presented prior to the dispens-

ing of a food pellet. During test sessions, a bar was inserted into the box, and each bar press was followed by one second of light. The 10 second duration group (10 seconds was the longest interval used) made significantly fewer responses than the others. The shape of the response-duration function is interesting in that response rate was greatest for the one second group, rather than for the .5 second group, which had the shortest delay between onset of light and of food. Although Bersh offers no reason for this, a discrimination viewpoint suggests an explanation of his data. Since all groups received the light for one second during test sessions, there is greatest similarity between training and test sessions for the one second training group. This, of course, assumes a rather fine temporal discrimination on the part of the rat. The test of this explanation would be an experiment in which duration of the light during test sessions is varied. A discrimination hypothesis leads to the prediction of a shift in the maxima point of the response-duration function with changes in the duration of the light during test sessions.

Jenkins (**22**) has also explored the role of the temporal interval between neutral and primary reinforcing stimulus presentation. His study differed from that of Bersh in two major respects. The test of secondary reinforcement was resistance to extinction of a bar-pressing response rather than learning, and the neutral stimulus, a buzzer, did not remain on during the interval preceding food presentation. Jenkins found that rate of response during extinction decreased monotonically as the interval between buzzer and food onset increased.

It appears that the interval between the onset of the neutral and the onset of the primary reinforcing stimulus does affect secondary reinforcement. Only subsequent research will reveal whether differences in the functions obtained by Bersh and Jenkins can be related to procedural differences.

Amount of Primary Reinforcement

The effect of varying the amount of primary reinforcement upon the strength of secondary reinforcement has been investigated by Lawson (**23**), D'Amato (**6**), and Hopkins (**21**). In the first of two experiments, Lawson used resistance to extinction of a runway response as a measure of secondary reinforcement. The secondary reinforcer was the end box in which food had been found previously. Although secondary reinforcement groups differed significantly from controls, there was no evidence of differential secondary reinforcement effects resulting from different amounts of reward during training. In a second experiment, an end box which had held food rewards during runway training was placed in one arm of a **U** maze. Again, no evidence was provided for significant effects of amount of primary reinforcement.

D'Amato provided for an apparently more sensitive measure of secondary reinforcement. Instead of comparing high and low reward groups, he gave all *S*s experience with high reward in one goal box, and with low reward in another. The test consisted of 15 choices between the two goal boxes (no food present) in a **T** maze. D'Amato found significant preferences for that goal box which had previously held the larger reward. D'Amato has suggested that the difference between his experimental results and those of Law-

son may be due to the fact that all of Lawson's *S*s encountered one goal box in which they had never before been rewarded. It seems more parsimonious to assume that the effect of amount of reward is so slight, that it can only be detected when the *S* is forced to choose between secondary reinforcers previously associated with different sized rewards.

Hopkins found no significant difference in single unit **T** maze learning for secondary reinforcement among five groups who differed in amount of primary reward received during training on a discrimination problem. Again, it is possible that the design, which was similar to Lawson's, did not permit a precise evaluation of the effects of amount of reward.

It appears that certain refinements of experimental design are necessary for the establishment of differential primary reinforcement effects. The D'Amato technique merits further application to determine whether differential effects can be shown consistently when each *S* chooses between high and low reward-associated secondary reinforcers. Investigators in this area might consider the problem of using D'Amato's design to investigate a wide range of amounts of reward, thereby establishing the shape of the function relating amount of reward and performance. The use of a multiple choice apparatus may be necessary for an efficient design.

The Role of Drive

Deprivation Schedules

Hall (**20**) investigated the relation between strength of drive during training and strength of secondary reinforcement during test sessions. Training consisted of learning a runway response to a particular goal box under either 6 or 22 hours' water deprivation. The rats then had to choose between the previously rewarded goal box and a previously unrewarded goal box in a **U** maze. Both groups were under 22-hour deprivation during this test of secondary reinforcement. There was no significant difference between the two groups in performance in the **U** maze.

Certain aspects of Hall's study deserve comment. No information is provided on the running times of the two groups during the training sessions. The absence of differences during the test sessions might have clearer implications if Hall had provided information about the effect of deprivation upon performance in the runway phase of the experiment. Data from other sources do not clarify the situation, since there is some disagreement in the literature regarding the effects of drive upon performance (**8**, **15**, **26**). Secondly, the investigation of only two points along the parameter leaves us with a rather incomplete picture of the relation between drive state during training and strength of secondary reinforcement. Finally, an evaluation of the strength of the secondary reinforcer used would have been facilitated if a control group had been run (possibly one which received a food reward in the **U** maze).

In a number of studies, drive state at the time of the secondary reinforcement test has been varied. Miles (**27**) conditioned rats in a Skinner box under 24 hours deprivation pairing each reinforcement with a light and a click. This stimulus complex was then presented during extinction sessions to six groups who had been deprived of food for either 0, 2½, 5, 10, 20 or 40 hours. The secondary reinforcement groups were superior to

comparably deprived control groups (in which light and click were absent during extinction). The response functions for both experimental and control groups were negatively accelerated, the distance between the curves increasing with hours of deprivation.

The Miles study is the only thorough parametric study of drive. However, Estes (**13**), Schlosberg and Pratt (**36**), Wike and Casey (**42**), and Seward and Levy (**38**) have all been concerned with the effect of satiation at the time of the test of secondary reinforcement strength.

Estes measured the learning of a bar press response in two groups of rats, one deprived of food for 23 hours and the other deprived for 6 hours. Both groups were "rewarded" only by the sound of the reinforcer, which had been previously associated with the presentation of water. The 6-hour group responded less than the 23-hour group, its rate being lower than it had previously been in an unrewarded free operant situation. Also, the 6-hour group did not differ significantly from a control group which had not had any secondary reinforcement training. It was concluded (*a*) that animals could be motivated by a secondary reinforcer when primary motivating conditions differed between training and test sessions and (*b*) that some type of strong primary motivation is necessary if the animal is to work for the secondary reinforcer.

Schlosberg and Pratt (**36**) used the sight and aroma of inaccessible food as a secondary reinforcer. Rats who were run in a single unit **T** maze while hungry learned to select the correct arm. Errors increased markedly whenever the rats were run while satiated.

Wike and Casey (**42**) rewarded the runway performance of satiated rats with food pellets. These animals performed better than controls who received nothing. There is a decided difference between the results of the Schlosberg and Pratt study, in which the subjects only saw and smelled the food, and the Wike and Casey study, in which the subjects were permitted to manipulate the food as well. It is conceivable that the Wike and Casey experiment involved a manipulatory drive, and that secondary reinforcement was not an effective factor. This view could be explored by running a satiated group which would find manipulable, inedible objects in the end-box.

The results of a study by Seward and Levy (**38**) are somewhat equivocal. The rats' preferences for either end-box position in a single unit **T** maze was determined. During the 9 days of training, half of the subjects received reinforcement (food and water) in the preferred end-box (Group P), half in the nonpreferred end-box (Group N). The animals were then run while satiated for 7 days. There was a significant difference in favor of Group P, in the proportion of choices of the originally preferred position. This difference did not increase over test days, indicating that while satiated rats can be motivated by secondary reinforcers, this motivation does not lead to improved learning over time.

Both resistance to extinction (**27**) and learning (**13, 36, 38, 42**) measures yield the conclusion that secondary reinforcement strength increases with increased drive state. There is still too little evidence to permit us to decide whether or not secondary reinforcement is effective with satiated subjects. Estes (**13**)

and Schlosberg and Pratt (**36**) have obtained negative results, while Wike and Casey (**42**) and Seward and Levy (**38**) have offered evidence to suggest that satiated subjects are influenced by secondary reinforcers. There appears to be a need for extensive, well-controlled research which will permit the statement of those experimental conditions under which satiated animals are motivated by secondary reinforcers.

Generalization of Drive

In a number of experiments, subjects have been tested for secondary reinforcement under motivating conditions other than those present during training. It has been previously pointed out that Estes (**13**) has shown that secondary reinforcers motivate behavior in the presence of a drive other than that present during training. Estes (**14**) also found transfer effects in two other experiments. However, the experimental group responded at a lower rate than a group learning under the original motivating conditions.

D'Amato (**7**) trained thirsty rats in a runway, rewarding them with water in a white cup in the goal box. They were then tested, while hungry, in a **T** maze, the correct response consisting of a turn to the former goal box. The correct turn was made significantly more often than chance expectancy, thus supporting the hypothesis that secondary reinforcement can be effective under motivating conditions other than those prevalent during training. Similar results were obtained in a second experiment, in which the rats were trained while hungry, and run in the **T** maze while thirsty.

Wike and Casey (**41**) found that thirsty, food-satiated rats ran faster to an end-box containing food pellets than did similarly motivated rats who ran to an empty end-box. This again suggests that secondary reinforcers are effective in the absence of the relevant drive. However, as in the previously considered study by these authors (**42**), there can be no assurance that the reinforcing factor was not the manipulability of the pellets.

In the first of two experiments, Reid and Slivinske (**34**) measured resistance to extinction of a bar pressing response which was consistently followed by a click. Two groups of rats had been trained to discriminate between food and water on the basis of a relevant drive. They were then deprived of food, and trained to bar press for food pellets, which were accompanied by a click. Half of these animals were then tested while hungry, the other half while thirsty. Another group of rats had no discrimination training, but were also trained to bar press for food pellets, and were also subdivided into two extinction groups. Assuming that generalization is a failure to discriminate, the authors predicted (*a*) that of the two hungry groups, the one with discrimination training would show greater resistance to extinction and (*b*) that of the two thirsty groups, the one without discrimination training would respond at a higher rate. In regard to the first hypothesis, no significant results were obtained. In regard to the second hypothesis, significant results, opposite to the predictions, were obtained. The authors concluded that secondary reinforcement, as they measured it, does not generalize from one drive to another.

It is apparent that the strength of

secondary reinforcement is greatest when measured under those drive conditions which were present during training. There is agreement, however, that secondary reinforcers are effective in the presence of an irrelevant drive, when the measure of their effectiveness is the learning of a new response. Reid and Slivinske, attempting to account for their failure to show transfer of secondary reinforcement have emphasized that they were measuring resistance to extinction. It is true that operational differences exist between the Estes studies on the one hand and the Reid and Slivinske study on the other. There is no obvious reason, however, for suspecting that the learning of a new response would occur in the presence of a different drive state, while the maintenance of an old one would not. It indeed appears quite reasonable to expect the reverse. It should also be noted that Webb (**40**) has shown that an irrelevant drive can be effective in maintaining a higher response rate during extinction.

The results of this area are further confounded by two factors, the difficulty of ensuring satiation, and the difficulty of equating drives. Estes (**14**) has pointed out that since thirst reduces food-intake, thirsty rats feeding ad libitum on food may not completely reduce their hunger. Osgood (**31**, p. 435) has suggested the use of drives other than thirst and hunger as a possible way of circumventing this difficulty. The second difficulty is more amenable to investigation. It might prove interesting to hold one drive constant during training and vary the extent of the irrelevant drive during test sessions, and to hold the irrelevant drive constant during the test, and vary the relevant drive during training.

Conclusions

The secondary reinforcement literature yields few conclusions which can be substantiated by a number of experiments, and which are not contradicted by other experiments. There is no consistent picture of the effects of schedules of reinforcement, amount of primary reinforcement, satiation or irrelevant drives. Both the discrimination hypothesis and the discriminative stimulus hypothesis must deal with experimental results which do not appear to be easily reconciled to these particular approaches. It is apparent that we have made only a tentative start towards an understanding of secondary reinforcement and the variables which affect its strength. There are a number of problems which future investigators might well consider.

Methodology

A number of writers have pointed out that two types of measures have been used in secondary reinforcement experiments, i.e. either resistance to extinction or learning of a new response. Reid and Slivinske (**34**) have suggested that different phenomena are involved and that the results of resistance to extinction studies may best be explained in terms of a discriminative stimulus hypothesis. Certainly the field is ripe for a systematic comparison of these two sets of operations. Can consistent differences be shown, over a number of parameters, when other variables (e.g. the neutral stimulus, the apparatus, the number of training trials) are held constant? A research program dealing with this

point might do much to yield a clear definition of secondary reinforcement, and of the operations necessary to establish it.

The measurement technique is only one of a number of variables which could account for the inconsistency of results in this area. Comparison of results is difficult when there is little standardization of apparatus, criterion, and controls. **T** mazes, runways, Skinner Boxes, and jumping stands have all been used. Subjects have been run for various numbers of trials, or for various periods of time, or until various periods of time have elapsed since the last response. Many of the studies are inadequately controlled in the sense that no comparison group (either a nonsecondary reinforcement group, or a primary reinforcement group) was run. The experimenter has the responsibility of indicating how powerful, or how slight, the effects of secondary reinforcement were at any level of the parameter under investigation. Saltzman's (**35**) use of a primary reinforcement group demonstrates that a powerful secondary reinforcement situation can be established in the laboratory. Miles' (**27**) investigations of drive and habit strength parameters permit comparison between secondary and nonsecondary reinforcement groups at all points on the parameters, and, in this respect, should serve as a model for future investigations. Unfortunately, there are a number of studies reviewed in this paper which lack such comparison groups.

The techniques used in data analysis have generally been unsophisticated. The *t* test has been used repeatedly in situations requiring the use of analysis of variance followed either by "post-mortem" tests (**39**) or, with ordered variables, orthogonal polynomials (**18**). Experiments in this area invariably involve measures taken over time or trials, and often over a number of points of a parameter. In many of these studies, information about the shape and rate of change of such functions could be obtained by the proper use of trend analysis techniques (**1, 18**). It is time that psychologists realized that the formulation of behavioral laws necessitates information about more than the differences in location of groups.

What is a Neutral Stimulus?

One of the most surprising aspects of research in this area is that little attention has been paid to the neutral stimulus . . . what it is and how it is related to the primary reinforcer in the learning situation. Experimenters have used buzzers, tones, clicks, and lights of various intensities as well as white and black end boxes as neutral stimuli. It is doubtful that these are all equally "neutral." It is quite possible that the particular stimulus chosen for pairing with the primary reinforcer is an important variable. There are recent studies (**17, 32**), for example, indicating that light functions as a reinforcer for rats, and that its effectiveness is related to its intensity. It would certainly be worthwhile to vary the intensity of the "neutral" stimulus, to ascertain whether any general relation exists between intensity and secondary reinforcement strength. The results of a study by Fink and Patton (**16**) are also suggestive. Rats learned a drinking response in the presence of a complex of light, sound, and touch. The removal of any of these caused a decrement in amount consumed, but the size of the decrement was a function of which component was re-

moved. The concept of a "neutral" stimulus therefore appears to be an unprofitable one. However, there are many stimuli whose reinforcer capacities change after pairing with a primary reinforcer. The extent of such increments in reinforcement capacity as they relate to the physical properties of the stimulus seems worthy of systematic investigation.

A distinction may be made between two techniques for establishing secondary reinforcers. The stimulus to be conditioned may be presented with, or immediately before, the primary reinforcer as in classical conditioning and as in the bulk of the studies reported here. On the other hand, the subject could be required to manipulate the stimulus in order to obtain the primary reinforcer, as in operant conditioning and as in Myers' (**29**) study with school children. In both situations the reinforcement capacity of some stimulus has been increased, but the training operations differ. One may ask whether either situation is generally more effective when other variables are held constant, and if so, whether this difference in effect is due to differences in proprioceptive feedback, the utilitarian aspect of the token, the period of time for which the stimulus is available on each trial, or some other factor. Here again is a problem which has received scant attention, but which might reveal information about those factors which facilitate the establishment of effective secondary reinforcers.

This paper has emphasized the evaluation of experimental conditions which may affect strength of secondary reinforcement and has ignored the role of secondary reinforcement as an explanatory concept in learning theory. Since the concept of secondary reinforcement has been widely regarded as a panacea for the reinforcement theorist's ills, and has therefore assumed a vital role in certain learning theories, this failure to relate the data to theory may perturb some. The author feels that secondary reinforcement is inadequately defined and inadequately demonstrated, that there is much disagreement about its relation to a number of variables, and that there are many gaps in our knowledge. Thus the use of secondary reinforcement as a foundation for any theory seems premature. The need, at this time, is for further and better research, rather than for more theory or defense of theory.

REFERENCES

1. ALEXANDER, H. A general test for trend. *Psychol. Bull.*, 1946, **43**, 533–557.
2. BERSH, P. J. The influence of two variables upon the establishment of a secondary reinforcer for operant responses. *J. exp. Psychol.*, 1951, **41**, 62–73.
3. BITTERMAN, M. E., FEDDERSEN, W. E., & TYLER, D. W. Secondary reinforcement and the discrimination hypothesis. *Amer. J. Psychol.*, 1953, **66**, 456–464.
4. BUGELSKI, R. Extinction with and without sub-goal reinforcement. *J. comp. Psychol.*, 1938, **26**, 121–134.
5. CLAYTON, F. L. Secondary reinforcement as a function of reinforcement scheduling. *Psychol. Reports*, 1956, **2**, 377–380.
6. D'AMATO, M. R. Secondary reinforcement and magnitude of primary reinforcement. *J. comp. physiol. Psychol.*, 1955, **48**, 378–380.
7. D'AMATO, M. R. Transfer of secondary reinforcement across the hunger and thirst drives. *J. exp. Psychol.*, 1955, **49**, 352–356.
8. DEVALOIS, R. L. The relation of different levels and kinds of motivation to variability of behavior. *J. exp. Psychol.*, 1954, **47**, 392–398.
9. DINSMOOR, J. A. A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *J. exp.*

Psychol., 1950, **40**, 458–472.

10. DINSMOOR, J. A. Resistance to extinction following periodic reinforcement in the presence of a discriminative stimulus. *J. comp. physiol. Psychol.*, 1952, **45**, 31–35.
11. DINSMOOR, J. A., KISH, G. B., & KELLER, F. S. A comparison of the effectiveness of regular and periodic secondary reinforcement. *J. gen. Psychol.*, 1953, **48**, 57–66.
12. ELAM, C. B., TYLER, T. W., & BITTERMAN, M. E. A further study of secondary reinforcement and the discrimination hypothesis. *J. comp. physiol. Psychol.*, 1954, **47**, 381–384.
13. ESTES, W. K. A study of motivating conditions necessary for secondary reinforcement. *J. exp. Psychol.*, 1949, **39**, 306–310.
14. ESTES, W. K. Generalization of secondary reinforcement from the primary drive. *J. comp. physiol. Psychol.*, 1949, **42**, 286–295.
15. FINAN, J. L. Quantitative studies in motivation. I. Strength of conditioning rats under varying degrees of hunger. *J. comp. Psychol.*, 1940, **29**, 119–134.
16. FINK, J. B., & PATTON, R. M. Decrement of a learned drinking response. *J. comp. physiol. Psychol.*, 1953, **46**, 23–27.
17. FORGAYS, D. G., & LEVIN, H. Learning as a function of sensory stimulation of various intensities. *Amer. Psychol.*, 1957, **12**, 411.
18. GRANT, D. A. Analysis-of-variance tests in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–154.
19. HALL, J. F. Studies in secondary reinforcement: I. Secondary reinforcement as a function of the frequency of primary reinforcement. *J. comp. physiol. Psychol.*, 1951, **44**, 246–251.
20. HALL, J. F. Studies in secondary reinforcement: II. Secondary reinforcement as a function of the strength of drive during primary reinforcement. *J. comp. physiol. Psychol.*, 1951, **44**, 462–466.
21. HOPKINS, C. O. Effectiveness of secondary reinforcing stimuli as a function of the quantity and quality of food reinforcement. *J. exp. Psychol.*, 1955, **50**, 339–342.
22. JENKINS, W. O. A temporal gradient of derived reinforcement. *Amer. J. Psychol.*, 1950, **63**, 237–243.
23. LAWSON, R. Amount of primary reward and strength of secondary reward. *J. exp. Psychol.*, 1953, **46**, 183–187.
24. MCCLELLAND, D. C., & MCGOWN, D. R. The effect of variable food reinforcement on the strength of a secondary reward. *J. comp. physiol. Psychol.*, 1953, **46**, 80–86.
25. MELCHING, W. H. The acquired reward value of an intermittently presented neutral stimulus. *J. comp. physiol. Psychol.*, 1954, **47**, 370–373.
26. MEYER, D. R. Food deprivation and discrimination reversal learning of monkeys, *J. exp. Psychol.*, 1951, **41**, 10–16.
27. MILES, R. C. The relative effectiveness of secondary reinforcers throughout deprivation and habit strength parameters. *J. comp. physiol. Psychol.*, 1956, **49**, 126–130.
28. MILLER, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology*. New York: Wiley, 1951. Pp. 435–472.
29. MYERS, N. A. Extinction of an operant response in children following partial and regular primary and secondary reinforcement procedures. Unpublished doctoral dissertation, Univer. of Wisconsin, 1957.
30. NOTTERMAN, J. M. A study of some relations among aperiodic reinforcement, discrimination training, and secondary reinforcement. *J. exp. Psychol.*, 1951, **41**, 161–169.
31. OSGOOD, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.
32. PREMACK, D., COLLIER, G., & ROBERTS, C. L. Frequency of light-contingent bar pressing as a function of the amount of deprivation for light. *Amer. Psychologist*, 1957, **12**, 411.
33. RATNER, S. C. Reinforcing and discriminative properties of the click in a Skinner box. *Psychol. Reports*, 1956, **2**, 332.
34. REID, L. S., & SLIVINSKE, A. J. A test for generalized secondary reinforcement during extinction under a different drive. *J. comp. physiol. Psychol.*, 1954, **47**, 306–310.
35. SALTZMAN, I. J. Maze learning in the absence of primary reinforcement: a study of secondary reinforcement. *J. comp. physiol. Psychol.*, 1949, **42**, 161–173.
36. SCHLOSBERG, H., & PRATT, C. H. The secondary reward value of food for hungry and satiated rats. *J. comp. physiol. Psychol.*, 1956, **49**, 149–152.

37. SCHOENFELD, W. N., ANTONITIS, J. J., & BERSH, P. J. A preliminary study of training conditions necessary for secondary reinforcement. *J. exp. Psychol.*, 1950, **40**, 40–45.
38. SEWARD, J. P., & LEVY, N. Choice-point behavior as a function of secondary reinforcement with relevant drives satiated. *J. comp. physiol. Psychol.*, 1953, **46**, 334–338.
39. STANLEY, J. C. Additional "post-mortem" tests of experimental comparisons. *Psychol. Bull.*, 1957, **54**, 128–130.
40. WEBB, W. B. The role of an irrelevant drive in response evocation in the white rat. *Amer. Psychologist*, 1947, **2**, 303.
41. WIKE, E. L., & CASEY, A. The secondary reinforcing value of food for thirsty animals. *J. comp. physiol. Psychol.*, 1954, **47**, 240–243.
42. WIKE, E. L., & CASEY, A. The secondary reward value of food for satiated animals. *J. comp. physiol. Psychol.*, 1954, **47**, 441–443.
43. WYCKOFF, L. B., SIDOWSKI, J., & CHAMBLISS, D. An experimental study of the relationship between secondary reinforcing and cue effects of a stimulus. *J. comp. physiol. Psychol.*, 1958, **51**, 103–109.

Received October 29, 1957.

PSYCHOLOGICAL BULLETIN
VOL. 55, NO. 5, 1958

CURIOSITY, EXPLORATORY DRIVE, AND STIMULUS SATIATION[1]

MURRAY GLANZER
American Institute for Research

In the 1930's, Dennis and his associates began the systematic study of exploratory behavior, centering their work on spontaneous alternation (**23, 24, 25, 75**). They determined the effect of such variables as repeated trials, end-arm preference, length of end-arm, number of choice points in tandem and number of alternatives on spontaneous alternation. These studies were followed by Heathers' study (**37**) which seemed to incorporate the behavior neatly within Hullian theory. The renewal of interest in the area was spurred by findings which showed that spontaneous alternation could not be explained by existing constructs and that it required new types of explanation (**30, 48**).

Since then there has been a growing interest in the behavior by which an organism increases its exposure to the various parts of its environment. The interest is, in part, a result of the regularities found in this type of behavior. These regularities are so striking that in several cases, the behavior forced itself upon the attention of investigators who were initially concerned with other variables (**57, 58, 60, 77**). The renewed interest is also, in part, a result of the decreased emphasis on orthodoxy in both concepts and areas of investigation. This relaxation followed the fading of the sharp theoretical boundaries of the thirties and forties.

[1] Based on a paper read at the Midwestern Psychological Association Meetings, 1957. I am indebted to Daniel E. Berlyne for his helpful comments on an earlier version of the paper.

The interest in exploratory behavior has manifested itself in well over 70 experimental and theoretical papers. The experiments concern the following situations:

1. Spontaneous alternation arrangements. A rat is permitted to make two successive choices of arms of a **T** maze. It usually displays a marked tendency to choose a different arm on its second trial than on its first. This occurs both with and without rewards.

2. Exploratory situations. An animal is placed in a complex environment such as a Dashiell maze. It will usually move so as to traverse a large number of units in this environment. (The spontaneous alternation arrangement may be viewed as a simplification of the exploratory situation.)

3. Reactivity tests. An animal is presented with a set of stimulus objects and the responses, i.e., approach, manipulation, are recorded. A change can then be introduced in the set of stimulus objects and the same observations made again. Responsiveness is greatest to new objects.

Work with these situations has yielded a number of clear findings. In general, the animal behaves so as to maximize contacts with new parts of the environment. Its responsiveness decreases systematically with continued exposure to the situation but recovers when the animal is absent from the environment. There have also been a number of learning situations used in which it has been demonstrated that animals will learn

in order to obtain access to new stimuli or exploratory situations.

Some of the work has generated, and in turn, been directed by three concepts: curiosity, exploratory drive, and stimulus satiation. The differences between these concepts, the new problems they lead to, some other approaches and the possibilities of broader formulation will be considered here.

Curiosity

The first concept, that of curiosity, was developed by Berlyne (**3**), as part of a program of extending Hullian principles to the area of perception. It is embodied in two postulates (**2**):

> When a novel stimulus affects an organism's receptors there will occur a drive-stimulus-producing response called curiosity.
>
> As a curiosity-arousing stimulus continues to affect an organism's receptors, curiosity will diminish. This diminution takes place under the Hullian model of extinction with both reactive inhibition and conditioned inhibition generated.

In other words, a novel stimulus leads to a response which, in turn, leads to a drive stimulus. With continued exposure there is extinction of the curiosity response.

Some findings out of the experimental work stimulated by this concept are the following:

1. That a place with numerous or complex stimuli is explored more than a place with few or simple stimuli. Rats entered an alcove in an exploratory box more frequently if it contained a wooden cube than if it was empty (**7**).

2. That the amount of exploration decreases sharply with continued exposure (**2, 7**).

3. That there is a permanent decrement in exploration after the first experience with a situation (**2, 7**). The number of entries into the alcove in the exploratory box declined in successive days.

There is further support for the first two points in the work of several experimenters (**13, 45, 72, 73**). There is some disagreement concerning the permanence of the effects of exposure to a set of stimuli or an environment. Montgomery does not find evidence for decrement from session to session (**47, 49**). Most of the evidence seems to support Berlyne's findings (**63, 72**). Berlyne has also done extensive work on human attention to change (**4**) and human curiosity (**5, 6**). For the latter area he employs a different set of constructs.

Exploratory Drive

The second concept is that of exploratory drive. This has been most extensively used by Montgomery. The concept has not been greatly elaborated. It is simply a drive that is elicited by novel stimuli. Montgomery also states that the strength of the tendency aroused by this drive is a decreasing function of time of exposure. Montgomery and his associates demonstrated the following points:

1. That spontaneous alternation is a function of stimulus rather than response characteristics. This was demonstrated in a cross-shaped maze with two starting boxes. When rats are started from opposite boxes in successive trials, they will alternate the end arm chosen even though this means repetition of the response, e.g., right turn (**48**).

2. That spontaneous alternation is independent of the amount of work done in the two alternatives. Rats run in a **Y** maze with rewards for pressing levers in the end arms showed no difference in amount of

alternation when the weights on the levers were varied (46).

3. That animals will learn to choose the alternative that leads to exploratory opportunities. Rats will learn to choose one arm of a **Y** maze or to make a black-white discrimination in order to gain access to a Dashiell maze (52, 55).

4. That the amount of exploratory behavior decreases regularly during one session in a maze and recovers after a period of absence from the maze (47, 49).

5. That decrement in exploratory behavior produced by one stimulus situation generalizes to other situations. Rats were permitted to explore successively **H** mazes that differed in luminance. The number of units traversed during a session on one maze depended on its similarity to the maze previously explored (49). The more similar the mazes, the less the amount of exploration on the succeeding maze.

6. That exploratory behavior is independent of opportunity for activity. Two groups of rats were housed in small cages. One group had access to an activity wheel, the other did not. No differences were found in the amount of maze exploration shown by the groups (51).

Montgomery and his co-workers also worked on the effects of hunger, thirst, fear and sensory deprivation upon exploratory behavior (50, 53, 54, 56, 79).

With respect to the findings on learning, Montgomery claims that the reinforcement is based not on drive reduction but on drive increase.[2] Berlyne and Slater (9) subsequently carried the investigation of learning based on exploratory reinforcement further. They found that the effective aspect of the reinforcer could be either the spaciousness of the end box or the complexity of stimulation presented by the end box.

[2] Drive increase reinforcement has been discussed at length by Leuba (42).

Stimulus Satiation

The concept of stimulus satiation was originally set up to furnish an alternative to a reactive inhibition analysis of spontaneous alternation. It is embodied in the following postulate.

When an organism observes a stimulus, a quantity of stimulus satiation is built up. This quantity reduces the responsiveness of the organism to the stimulus. The longer the stimulus is present, the greater the amount of stimulus satiation built up. In the absence of the stimulus, this quantity dissipates. It is also postulated that stimulus satiation generalizes to similar stimuli (30).

The concept contains elements similar to both the reactive inhibition and the conditioned inhibition constructs of Hull. It is similar to reactive inhibition in the way in which it builds up and dissipates. It is similar to conditioned inhibition in that it is tied to environmental stimuli. Rothkopf and Zeaman[3] (59), and Broadbent (11) have presented similar ideas.

On the basis of the postulate a number of assertions were made and found to be supported by previous or subsequent experimental work. Some of the results are the following:

1. Spontaneous alternation was demonstrated to be a function of stimulus factors, rather than response factors, (namely, reactive inhibition) (29). This was accomplished by means of the cross-shaped maze described above.

2. The effect of time interval be-

[3] Zeaman and his associates have carried out an intensive program of investigation in this area (59, 76, 77).

tween spontaneous alternation trials was shown to depend on the place in which the animal was detained (**29**). Delay between trials usually decreases the amount of spontaneous alternation. When, however, rats were detained in the end arm they chose on the first trial, the amount of alternation was increased.

3. Walker, et al. separated the effects of extramaze and intramaze stimuli in spontaneous alternation, showing that both play a role (**68**). They used a maze that could be rotated about the choice point. When the maze was rotated between trials the animals were forced to alternate on the basis of either the intramaze or the extramaze stimuli.

4. These investigators also reevaluated the role of response factors. They demonstrated that if a response eventuated in differential stimulation it could determine spontaneous alternation (**70**). This was done by the use of a maze with banked pathways that accentuated the differences in the alternative responses available to the animals.

5. Dember found that a varying alternative was preferred to a constant one (**21**) and that the greater the difference between variations, the greater the preference (**22**). In the experiments a **T** or **Y** maze with glass doors blocking entry to the end arms was used. The rats were exposed to both end arms by being permitted to explore the choice point. The glass doors were then removed allowing the animal to make its choice of one of the alternatives. In the first experiment cited above the arms differed in color, e.g., one black and one white, during the exposure period. Immediately before the choice was permitted one of the alternatives was changed so that it was the same as the other. The animals prefer the changed alternative. In the second experiment the two alternatives also differed in color, e.g., one black and one gray, and then were both changed to another color, e.g., both to white. The animals prefer the arm showing the larger change (black to white).

6. Walker (**67**) demonstrated a correlation between the discriminability of alternatives as measured in learning situations and the amount of spontaneous alternation they elicited. A cross-shaped, rotating maze was used that permitted the isolation of the effects of intra-maze, extramaze and response-derived cues. It was found that those cues that influenced spontaneous alternation most strongly also were most effective as a basis for learning. For example, intra-maze cues produced more spontaneous alternation than extra-maze cues. If reward was associated with intra-maze cues it produced faster learning than if it was associated with extra-maze cues. (The maze was rotated and the starting arm changed from trial to trial. In the first case the animal had to learn to choose a given arm, in the second case he had to learn to choose the arm on a given side of the room.)

There have been a number of new developments in this area. Denny (**26**) has carried the concept into the area of learning effects and has presented evidence that rats will learn to choose the less frequently experienced of two alternatives. Walker's work has led to his developing certain new views concerning neural action and the effect of reward (**66, 67**). A number of questions were raised concerning the stimulus satiation concept. One question concerned the manner in which an organism had to be exposed to an alternative in order for stimulus satiation effects to appear.

1. Walker, et al., (69) on the basis of negative findings with a technique involving preliminary exposure of the animals to one alternative, suggested that stimulus satiation occurs only when a choice is made.

2. Kivy, et al. (41) argued that active choice was not the prerequisite but that exposure to the alternatives in the context of the choice point was. These investigators allowed the animals to preview two similar alternatives through glass doors at the choice point. One alternative was then changed, the glass doors removed, and the animals permitted to make a free choice. They chose the new, dissimilar alternative.

Later findings, however, have indicated that neither the active choice nor the choice point context restriction may be necessary.

1. Sutherland (61) found that when animals were prefed in one of two end boxes of a T maze and were then permitted to make a free choice, they chose the arm leading to the other end box.

2. In another experiment (31), using procedures similar to Walker, et al. (69), the animals chose the alternative leading to the unexperienced end box.

Another question concerns the aspect of the stimulus that is effective in producing stimulus satiation. Dember (21) argues that novelty or stimulus change is the crucial characteristic. Berlyne (7, 9) has carried out more extensive investigation of the stimulus characteristics effective in eliciting exploratory behavior. Berlyne has also presented criticism of the concept on the basis of findings with human subjects (8).

Manipulation Motive

Harlow has presented the same elicited-drive notion found in Montgomery and Berlyne's concepts in his discussions of the manipulation motive. A number of experiments by him and his associates have been concerned with showing the persistence and reliability of the behavior involved and that it can be used as a basis for learning (20, 32, 33, 35, 36).

Harlow indicated some time ago that the two drives, manipulation motive and exploration motive, are distinct on the basis of the considerable persistence of the former (35). Butler's more recent work has, however, shown the same persistence for behavior of this exploratory type (12, 16, 17). The basis for distinction between the two, therefore, no longer exists.

It may be suggested that both types of behavior be subsumed under one of the concepts introduced thus far. The advantage of doing this is that it suggests a number of aspects of manipulatory behavior that might display regularities similar to those found in exploratory behavior. Two predictions suggest themselves if this unified explanation is used and the Harlow puzzles are viewed as giving the subject stimulus changes as he unlocks them:

1. Pre-exposure to the puzzles, particularly in their open state, will result in a lower rate of opening the puzzles.

2. Puzzles that yield maximum stimulus change will be learned faster than those that present minimal stimulus change. For example, if Harlow's puzzles were painted so that the hasp was blue on one side and yellow on the other, then opening the hasp would produce a marked change. It is predicted then that the monkey would learn faster than if both sides of the hasp were the same color.

If these predictions proved accu-

rate, further predictions could be taken from the stimulus satiation postulate, for example, and translated into the Harlow puzzle situation. In the case of stimulus satiation, the predictions would concern the effect of such variables as number of parts of the puzzle, and presence or absence of the puzzle during intertrial intervals.

Empirical Investigations

There have been several investigators carrying out studies with little allegiance to any theoretical constructs. One of their major accomplishments has been the extension of the findings concerning exploratory behavior to other species.

1. Butler (**12**, **13**, **14**, **15**, **16**, **17**) has demonstrated, with rhesus monkeys, that exploratory activity is highly persistent, that it recovers from satiation, and can be used as a reinforcer in instrumental response learning and discrimination learning. Butler has developed, for his work, a variant of the Skinner box in which the operant is window-pushing.

2. Thompson and his collaborators (**62**, **63**, **64**, **65**) have carried out investigations on both rats and dogs. The work with dogs has indicated the usual decline in exploratory activity during a session. It also shows differences in amount of exploration as a function of age and amount of sensory deprivation.

3. Welker (**71**, **72**, **73**) has used chimpanzee subjects and found the same declines in exploratory behavior within a session, recovery after a rest period and an overall decline in repeated sessions. He also found that complex, changing objects were preferred and that amount of exploratory behavior was a function of age.

Differences Between the Three Concepts

The three concepts—curiosity, exploratory drive, and stimulus satiation are very similar. The similarity may be seen in the fact that almost all of the findings mentioned can be derived using any one of the three concepts. All three involve initial changes in behavior when a new stimulus is presented. All three involve decrements in this behavior with continued exposure.

There are some differences and these are of two types: First, the gross difference in the words employed; second, the differences in implication. The difference in the wording has a heuristic effect. The associations aroused by stimulus satiation are somewhat different from those aroused by exploratory drive. From the first come experimental problems on stimulus characteristics, e.g., extramaze stimuli versus intramaze stimuli. From the second come problems concerned with learning based on exploratory drive. The wording difference is not so great as to prevent experiments on stimulus characteristics in the name of exploratory drive (**49**) or experiments on learning in the name of stimulus satiation (**26**, **74**). Each of the concepts has, however, characteristics that distinguish it from the others.

The unique characteristic of the exploratory drive is that, modeled after a drive like hunger, it does not allow for long term effects of repeated exposure to a situation. Several times Montgomery underlined the fact that there was complete recovery of exploratory activity from one daily session to the next (**47**, **49**). Other findings as indicated above, do not concur on this point (**2**, **7**, **72**). Berlyne's curiosity postulate covers

these other findings by the presence of conditioned inhibition of the curiosity response. Stimulus satiation can account for such effects by keeping the asymptote, to which it declines, above zero.

The distinguishing characteristic of both exploratory drive and curiosity as compared with stimulus satiation is that the former work by increasing responsiveness and activity level to new stimuli while the latter works by decreasing responsiveness and activity level to previously experienced stimuli. This has some implication for learning based on these concepts as will be pointed out later.

These in general are the differences between the concepts. Curiosity as opposed to exploratory drive involves conditioned inhibition of drive. Stimulus satiation as opposed to both uses decremental rather than drive factors.

Problems

Of greater interest, perhaps, (than the differences between the three concepts) are some of the new problems and techniques that have arisen in this area. These fall into the following four areas.

1. Learning
2. Drive
3. Stimulus effects
4. Fear of novel stimuli

Learning

There have been numerous learning experiments carried out in this area. In general, the subjects learn to choose the alternative which leads to stimulus changes or opportunity for exploratory behavior. In experiments such as Denny's, the subjects learn to choose the alternative to which they had run less frequently. In other experiments the subjects learn to choose the alternative that leads to a Dashiell maze or allows them to peek into another room. Even the Skinner box has yielded pertinent data. After a tradition of experiments in which rodents were demonstrated to learn in order to turn off lights, Kish, Marx, and others (**40, 44**) discovered that they will also learn in order to turn them on. The key factor seems to be variation of stimulation rather than light-avoidance or light-approach.

Certain issues may be raised concerning the mechanism by which such learning takes place. One is related to whether the animal is learning to approach one alternative or learning not to approach the other. This question can be given operational meaning in terms of the latency data obtainable in these learning experiments. If there are decreased running times or latencies as the trials continue, the mechanism would seem to be learning to approach a positive alternative. If, however, the running times or latencies increase then, the mechanism is more of the type of relatively reduced responsiveness to one alternative. This is distantly related to the question of whether a drive-like or satiation-like mechanism is involved. At the present time the data on this point are not clear. This is as far as the issue can be pushed at the present time. No comparison of theories is possible until further spelling out of the various mechanisms occurs. For example, what is the relative speed with which the curiosity drive extinguishes as compared with learning based on it?

Drive

Another major concern is with the effect of drives such as hunger upon exploratory behavior. The results again have been somewhat conflict-

ing. Dashiell (**19**) in 1925 found that hungry animals explored more than satiated animals. Montgomery (**50**) found that they explored less, Adlerstein and Fehrer (**1**) that they explored more, and Thompson (**62**) found no significant differences.

The conflict of results is not surprising. First, on theoretical grounds the introduction of a drive has been postulated to have two effects.

1. It energizes more strongly whatever behavior is ongoing.
2. It introduces drive stimuli.

With respect to exploratory behavior these would have opposed effects. The energizing aspect of drive should give rise to more exploratory behavior. However, additional internal stimuli would reduce the relative importance of the external stimuli that the animal responds to in its exploratory behavior. Therefore, there should be less exploratory behavior. Adding to the theoretical problem is the fact that the experiments were done in mazes that permitted a multiplicity of choices. Using a complex maze for the measurement of these effects is something like using a group discussion situation to collect psychophysical data. A number of side effects can enter in to completely confound the results. It has been pointed out (**7**) that one cannot easily tell in the free exploration situation whether the animal is standing still because he is not curious about what is around him or because he is very curious. Fehrer (**28**) and Zimbardo and Miller (**78**) have both recently simplified the exploratory situation by manipulating familiarity of parts of the test situation and have come up with somewhat clearer results favoring the original Dashiell assertion. Fehrer, for example, measured the amount of time spent in a new exploratory box when one group ran from a familiar starting box and the other from a new, unfamiliar starting box. The first group spent more time in the exploratory box. This difference was greater when the animals were hungry than when the animals were satiated. Further support for the Dashiell assertion may be found in the work of Campbell and Sheffield (**18**) on the effect of hunger on responsiveness to environmental changes.

The conflicting points on a theoretical basis still remain. A prediction could be made that further experimental work will reveal that the results obtained depend on how the two opposed effects of drive (energizing and stimulus-producing) interact with aspects of the experimental situation. For example, it could be predicted that drive state would interact significantly with the discriminability of the alternatives presented. Furthermore, using experimental procedures that vary the stimulus-producing effects of drive (e.g., using animals with varying degrees of adaptation to cycles of food deprivation or using animals trained in varying degrees to respond to drive-produced stimuli) results in either direction could be obtained.

Stimulus Effects

One of the most exciting characteristics of the work is the fact that it has led to the development of new techniques for answering the primordial question that gave rise to a major part of comparative psychology and the psychology of animal learning. The question is "How does the animal see his world?" The first major inprovement in collecting data to answer this question was the Lashley jumping stand. The second was the equivalent stimulus method of Klüver. These techniques although

considerably better than what preceded them are all tedious and time-consuming and animal-consuming. They are animal-consuming because the effect of learning one discrimination upon later discriminations poses problems that can be avoided only with the use of new subjects. The costliness of getting information in this manner probably led to the fact that work in this area has fallen far short of the promise implicit in the early work of Lashley and Klüver. One result of this failure is seen in the absorption with the response side of behavior. Much psychological theory is dedicated to the proposition that it is impossible to know very much about the stimuli the animal is responding to.

Out of the recent work on curiosity and stimulus satiation has come a simple technique that gives the experimenter easy access to the perceptual world of the animal. To find out how different A and B are to a subject, the experimenter simply allows him to satiate to one and then presents the second. The speed and frequency of approach is the measure of the difference of the two for the subject. This technique has been embodied in both the Kivy-Dember glass choice point technique (**21**, **22**, **41**) and the Berlyne-Thompson exploratory cubicle technique (**7**, **65**). Information about stimulus differences that in the past had been obtained only after weeks of training can in this manner be determined in a day or two. Furthermore, the subject is available for further use either in continued work of this nature or for other types of experimentation since the effect of this type of exposure seems to be relatively transitory and with slight transfer effect.

Fear of Novel Stimuli

One of the major problems that requires handling concerns the fact that novel stimuli may give rise to avoidance and a number of other fear-related responses as well as approach. This point has been discussed by a number of investigators (**38**, **39**, **53**). Furthermore, the data of a number of the studies of exploratory behavior reflect these fear effects. For example, some of the data do not display the usual simple decay function for amount of exploratory activity within a continuous session. In many cases there is an initial rise that may be ascribed to reduction of fear. And anyone who has handled animals in preparation for experimental work is well aware of the strength of these fear reactions to new situations. Any complete explanation must cover these effects.

The two drive formulations are particularly incomplete in this respect. Montgomery has postulated a twin fear-of-novel-stimuli drive. But this raises the question of predicting when fear and when exploratory drive should predominate.

The stimulus satiation concept skirts this problem by leaving the response unspecified. The postulate merely says that novel stimuli bring out more strongly than familiar stimuli whatever response the animal would make at that time to the situation. Approach or avoidance tendencies may be conceived of as initially determined by such operations as handling of the animal. This is, however, not quite satisfactory. In order to handle these effects, it is necessary to move to a more general system. One possibility will be considered next in the discussion of general problems raised by the work considered above.

General Issues

One of the general issues that has been raised concerns the ancient na-

tivist-empiricist controversy. There is, however, not likely to be much profit in disputes over whether the behavior is originally or primarily learned or unlearned. The important thing here is that some behavioral regularities have been found. Anything that furthers knowledge of the factors that control them or are associated with them deserves attention.

Harlow's arguments that his exteroceptive drives must be unlearned and primary are not completely convincing. For example, he has emphasized the fact that introducing food rewards disrupts the manipulation behavior and reasons from this that secondary reward cannot be involved. This argument must be questioned. The manipulation puzzles presented to the monkeys are detour or Umweg problems. Direct attack on the part of the puzzle holding the goal object prevents solution. There are indications in previous work that raising motivation and incentive levels (**10**, **43**) has an interfering effect on solution of problems of this type. Harlow also rejects the learned basis of this manipulation motive since he claims manipulation has not been rewarded in the past. His views are supported by his recent findings with Blazek and McClearn (**34**) on the early appearance of manipulatory behavior in infant rhesus monkeys. The findings are, however, far from complete support for the argument.

Estes and Schoeffler (**27**) on the other hand, have recently suggested some experimental work to prove that spontaneous alternation can be learned. It is doubtful whether such experiments would settle the issue. The fact that the eyeblink can be conditioned does not mean that people blink their eyes because they have been conditioned to do so. Whether the behavior is learned or innate is a genetic question and can be settled, if at all, only by a detailed genetic analysis of the behavior.

Is there a more general system in which the concepts discussed here might be incorporated? One which immediately suggests itself is a historical learning model. This seems, however, to lead into the empiricist-nativist controversy rather than into the elucidation of new regularities in behavior. A somewhat different system may satisfy the requirements implied above.

This system would stem from the following observations:

1. Animals generally seek out new stimuli.

2. Sometimes, new stimuli cause fright and freezing instead of approach and exploration.

3. The animal's reaction is affected by its previous experience, e.g., handling, prolonged isolation, etc.

The first point has been borne out experimentally and systematized by the concepts discussed above. The second point has been noted a number of times. The third point has received some study but the facts involved are, as yet, unclear (**56**, **63**, **79**). Clarification will probably occur when the exploratory test situations are simplified as has been done in the case of the recent studies on the effects of hunger on exploratory behavior (**28**, **78**).

The general system proposed here is the following. The organism is viewed as an information processing system that requires certain amounts of information per unit time. If the organism's environment does not present this amount then the organism is activated either to increase the amount that is presented by increased locomotion or to decrease the amount by avoidance, freezing, seizure, etc. No way has been presented,

however, for determining whether the organism will approach or avoid new stimuli. The following postulate is, therefore, added: That the organism's information requirements are set by its past experience. An organism that has had a high flow of information directed at it in the past would have a high requirement or standard. An organism that has lived in an impoverished informational environment would have a low requirement or standard. The organism would respond in terms of the difference between its individual standard and the amount of information furnished by the situation.

These ideas may be summarized in the following statement: The increase or decrease of activity with respect to parts of the environment is a function of the difference between the average amount of information the individual is accustomed to and the current rate of flow of information from the environment. The statement can be summarized as

$$\frac{dA}{dt} = f\left(\frac{I}{t} - \frac{dI}{dt}\right)$$

where A is amount of activity, I is amount of information processed during the organism's life history and t is time measured from the birth of the organism.

There are, of course, other factors that determine rate of change of activity. For example, the equation should contain additional terms to take into account limits on the activity of the organism. Another equation may also be added describing the relation of dI/dt to A.

From even this sketchy formulation, a number of new implications can be seen. These are in the following areas:

1. The differential effects of early and late experience. New information can have a much greater effect on changing the organism's standard (I/t) and, therefore, its behavior, early in life than late in life.

2. The effects of aging. There are several factors that keep an organism in a limited area (e.g., fatigue, home territory, food supply). Therefore, as it grows older, it will receive less and less new information. This in the long run will affect its information requirement or standard. There are some data relevant to this point (**63**, **72**).

In summary, three concepts employed in this area have been compared; some of the new problems and techniques that they have made prominent have been mentioned, and possibilities for a broader framework for the area have been suggested. The final word will come, of course, from the laboratory.

REFERENCES

1. ADLERSTEIN, A., & FEHRER, E. The effect of food deprivation on exploratory behavior in a complex maze. *J. comp. physiol. Psychol.*, 1955, **48**, 250–253.
2. BERLYNE, D. E. Novelty and curiosity as determinants of exploratory behavior. *Brit. J. Psychol.*, 1950, **41**, 68–80.
3. BERLYNE, D. E. Attention, perception and behavior theory. *Psychol. Rev.*, 1951, **58**, 137–146.
4. BERLYNE, D. E. Attention to change. *Brit. J. Psychol.*, 1951, **42**, 269–278.
5. BERLYNE, D. E. A theory of human curiosity. *Brit. J. Psychol.*, 1954, **45**, 180–191.
6. BERLYNE, D. E. An experimental study of human curiosity. *Brit. J. Psychol.*, 1954, **45**, 256–265.
7. BERLYNE, D. E. The arousal and satiation of perceptual curiosity in the rat. *J. comp. physiol. Psychol.*, 1955, **48**, 238–246.
8. BERLYNE, D. E. Attention to change, conditioned inhibition (sI_R), and stimu

lus satiation. *Brit. J. Psychol.*, 1957, **48**, 138–140.

9. BERLYNE, D. E., & SLATER, J. Perceptual curiosity, exploratory behavior and maze learning. *J. comp. physiol. Psychol.*, 1957, **50**, 228–232.
10. BIRCH, H. G. The role of motivational factors in insightful problem-solving. *J. comp. Psychol.*, 1945, **38**, 295–317.
11. BROADBENT, D. E. Classical conditioning and human watch-keeping. *Psychol. Rev.*, 1953, **60**, 331–339.
12. BUTLER, R. A. Discrimination learning by rhesus monkeys to visual-exploration motivation. *J. comp. physiol. Psychol.*, 1953, **46**, 95–98.
13. BUTLER, R. A. Incentive conditions which influence visual exploration. *J. exp. Psychol.*, 1954, **48**, 19–23.
14. BUTLER, R. A. The effect of deprivation of visual incentives on visual exploration motivation in monkeys. *J. comp. physiol. Psychol.*, 1957, **50**, 177–179.
15. BUTLER, R. A. Discrimination learning by rhesus monkeys to auditory incentives. *J. comp. physiol. Psychol.*, 1957, **50**, 239–241.
16. BUTLER, R. A., & ALEXANDER, H. M. Daily patterns of visual exploratory behavior in the monkey. *J. comp. physiol. Psychol.*, 1955, **48**, 247–249.
17. BUTLER, R. A., & HARLOW, H. F. Persistence of visual exploration in monkeys. *J. comp. physiol. Psychol.*, 1954, **47**, 258–263.
18. CAMPBELL, B. A., & SHEFFIELD, F. D. Relation of random activity to food deprivation. *J. comp. physiol. Psychol.*, 1953, **46**, 320–322.
19. DASHIELL, J. F. A quantitative demonstration of animal drive. *J. comp. Psychol.*, 1925, **5**, 205–208.
20. DAVIS, R. T., SETTLAGE, P. H., & HARLOW, H. F. Performance of normal and brain-operated monkeys on mechanical puzzles, with and without food incentive. *J. genet. Psychol.*, 1950, **77**, 305–311.
21. DEMBER, W. N. Response by the rat to environment change. *J. comp. physiol. Psychol.*, 1956, **49**, 93–95.
22. DEMBER, W. N., & MILLBROOK, B. A. Free-choice by the rat of the greater of two brightness changes. *Psychol. Reports*, 1956, **2**, 465–467.
23. DENNIS, W. A comparison of the rat's first and second exploration of a maze unit. *Amer. J. Psychol.*, 1935, **47**, 488–490.
24. DENNIS, W. Spontaneous alternation in rats as an indicator of the persistence of stimulus effects. *J. comp. Psychol.*, 1939, **28**, 305–312.
25. DENNIS, W., & SOLLENBERGER, R. J. Negative adaptation in the maze exploration of albino rats. *J. comp. Psychol.*, 1934, **18**, 197–206.
26. DENNY, M. R. Learning through stimulus satiation. *J. exp. Psychol.*, 1957, **54**, 62–64.
27. ESTES, W. K., & SCHOEFFLER, M. S. Analysis of variables influencing alternation after forced trials. *J. comp. physiol. Psychol.*, 1955, **48**, 357–362.
28. FEHRER, E. The effects of hunger and familiarity of locale on exploration. *J. comp. physiol. Psychol.*, 1956, **49**, 549–552.
29. GLANZER, M. The role of stimulus satiation in spontaneous alternation. *J. exp. Psychol.*, 1953, **45**, 387–393.
30. GLANZER, M. Stimulus satiation: An explanation of spontaneous alternation and related phenomena. *Psychol. Rev.*, 1953, **60**, 257–268.
31. GLANZER, M. Stimulus satiation in situations without choice. *J. comp. physiol. Psychol.*, 1958, **51**, 332–335.
32. HARLOW, H. F. Learning and satiation of response in intrinsically motivated complex puzzle performance by monkeys. *J. comp. physiol. Psychol.*, 1950, **43**, 289–294.
33. HARLOW, H. F. Motivation as a factor in the acquisition of new responses. In J. S. Brown, et al. (Ed.), *Current theory and research in motivation*, Lincoln: Univer. Nebraska Press, 1953, 24–29.
34. HARLOW, H. F., BLAZEK, NANCY C., & MCCLEARN, G. E. Manipulatory motivation in the infant rhesus monkey. *J. comp. physiol. Psychol.*, 1956, **49**, 444–448.
35. HARLOW, H. F., HARLOW, M. K., & MEYER, D. R. Learning motivated by a manipulation drive. *J. exp. Psychol.*, 1950, **40**, 228–234.
36. HARLOW, H. F., & MCCLEARN, G. E. Object discrimination learned by monkeys on the basis of manipulation motives. *J. comp. physiol. Psychol.*, 1954, **47**, 73–76.
37. HEATHERS, G. L. The avoidance of repetition of a maze reaction in the rat as a function of the time interval between trials. *J. Psychol.*, 1940, **10**, 359–380.
38. HEBB, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, **62**, 243–254.

39. HEBB, D. O. *The organization of behavior.* New York: Wiley, 1949.
40. KISH, G. B. Learning when the onset of illumination is used as reinforcing stimulus. *J. comp. physiol. Psychol.*, 1955, **48**, 261–264.
41. KIVY, P. N., EARL, R. W., & WALKER, E. L. Stimulus context and satiation. *J. comp. physiol. Psychol.*, 1956, **49**, 90–92.
42. LEUBA, C. Toward some integration of learning theories: the concept of optimal stimulation. *Psychol. Reports*, 1955, **1**, 27–33.
43. LEWIN, K. *A dynamic theory of personality.* New York: McGraw-Hill, 1935.
44. MARX, M. H., HENDERSON, R. L., & ROBERTS, C. L. Positive reinforcement of the barpressing response by a light stimulus following dark operant pretests with no after effect. *J. comp. physiol. Psychol.*, 1955, **48**, 73–76.
45. MONTGOMERY, K. C. The relation between exploratory behavior and spontaneous alternation in the white rat. *J. comp. physiol. Psychol.*, 1951, **44**, 582–589.
46. MONTGOMERY, K. C. "Spontaneous alternation" as a function of time between trials and amount of work. *J. exp. Psychol.*, 1951, **42**, 82–93.
47. MONTGOMERY, K. C. Exploratory behavior and its relation to spontaneous alternation in a series of maze exposures. *J. comp. physiol. Psychol.*, 1952, **45**, 50–57.
48. MONTGOMERY, K. C. A test of two explanations of spontaneous alternation. *J. comp. physiol. Psychol.*, 1952, **45**, 287–293.
49. MONTGOMERY, K. C. Exploratory behavior as a function of "similarity" of stimulus situation. *J. comp. physiol. Psychol.*, 1953, **46**, 129–133.
50. MONTGOMERY, K. C. The effect of hunger and thirst drives upon exploratory behavior. *J. comp. physiol. Psychol.*, 1953, **46**, 315–319.
51. MONTGOMERY, K. C. The effect of activity deprivation upon exploratory behavior. *J. comp. physiol. Psychol.*, 1953, **46**, 438–441.
52. MONTGOMERY, K. C. The role of exploratory drive in learning. *J. comp. physiol. Psychol.*, 1954, **47**, 60–64.
53. MONTGOMERY, K. C. The relation between fear induced by novel stimulation and exploratory behavior. *J. comp. physiol. Psychol.*, 1955, **48**, 254–260.
54. MONTGOMERY, K. C., & MONKMAN, J. A. The relation between fear and exploratory behavior. *J. comp. physiol. Psychol.*, 1955, **48**, 132–136.
55. MONTGOMERY, K. C., & SEGALL, M. Discrimination learning based upon the exploratory drive. *J. comp. physiol. Psychol.*, 1955, **48**, 225–228.
56. MONTGOMERY, K. C., & ZIMBARDO, P. G. The effect of sensory and behavioral deprivation upon exploratory behavior in the rat. *Percept. Mot. Skills*, 1957, **7**, 223–229.
57. MOTE, F. A., JR., & FINGER, F. W. Exploratory drive and secondary reinforcement in the acquisition and extinction of a simple running response. *J. exp. Psychol.*, 1942, **31**, 57–69.
58. MYERS, A. K., & MILLER, N. E. Failure to find a learned drive based on hunger; evidence for learning motivated by "exploration." *J. comp. physiol. Psychol.*, 1954, **47**, 428–436.
59. ROTHKOPF, E. Z., & ZEAMAN, D. Some stimulus controls of alternation behavior. *J. Psychol.*, 1952, **34**, 235–255.
60. SCHOENFELD, W. N., ANTONITIS, J. J., & BERSH, P. J. Unconditioned response rate of the white rat in a bar pressing apparatus. *J. comp. physiol. Psychol.*, 1950, **43**, 41–48.
61. SUTHERLAND, N. S. Spontaneous alternation and stimulus avoidance. *J. comp. physiol. Psychol.*, 1957, **50**, 358–362.
62. THOMPSON, W. R. Exploratory behavior as a function of hunger in "bright" and "dull" rats. *J. comp. physiol. Psychol.*, 1953, **46**, 323–326.
63. THOMPSON, W. R., & HERON, W. The effects of early restriction on activity in dogs. *J. comp. physiol. Psychol.*, 1954, **47**, 77–82.
64. THOMPSON, W. R., & KAHN, A. Retroaction effects in the exploratory activity of "bright" and "dull" rats. *Canad. J. Psychol.*, 1955, **9**, 173–182.
65. THOMPSON, W. R., & SOLOMON, L. M. Spontaneous pattern discrimination in the rat. *J. comp. physiol. Psychol.*, 1954, **47**, 104–107.
66. WALKER, E. L. The duration and course of the reaction decrement and the influence of reward. *J. comp. physiol. Psychol.*, 1956, **49**, 167–176.
67. WALKER, E. L. Action decrement and its relation to learning. *Psychol. Rev.*, 1958, **65**, 129–142.
68. WALKER, E. L., DEMBER, W. N., EARL, R. W., & KAROLY, A. J. Choice alternation: I. Stimulus vs. place vs. re-

sponse. *J. comp. physiol. Psychol.*, 1955, **48**, 19–23.

69. WALKER, E. L., DEMBER, W. N., EARL, R. W., FLIEGE, S. E., & KAROLY, A. J. Choice alternation: II. Exposure to stimulus or stimulus and place without choice. *J. comp. physiol. Psychol.*, 1955, **48**, 24–28.
70. WALKER, E. L., DEMBER, W. N., EARL, R. W., FAWL, C. L., & KAROLY, A. J. Choice alternation: III. Response intensity vs. response discriminability. *J. comp. physiol. Psychol.*, 1955, **48**, 80–85.
71. WELKER, W. I. Effects of age and experience on play and exploration of young chimpanzees. *J. comp. physiol. Psychol.*, 1956, **49**, 223–226.
72. WELKER, W. I. Some determinants of play and exploration in chimpanzees. *J. comp. physiol. Psychol.*, 1956, **49**, 84–89.
73. WELKER, W. I. Variability of play and exploratory behavior in chimpanzees. *J. comp. physiol. Psychol.*, 1956, **49**, 181–185.
74. WELLS, R. H. Learning through stimulus satiation: one trial a day. Unpublished doctoral dissertation, Michigan State Univer., 1956.
75. WINGFIELD, R. C., & DENNIS, W. The dependence of the rat's choice of pathways upon the length of the daily trial series. *J. comp. Psychol.*, 1934, **18**, 135–147.
76. ZEAMAN, D., & ANGELL, D. A spatial gradient of alternation tendency. *J. comp. physiol. Psychol.*, 1953, **46**, 383–386.
77. ZEAMAN, D., & HOUSE, B. J. The growth and decay of reactive inhibition as measured by alternation behavior. *J. exp. Psychol.*, 1951, **41**, 177–186.
78. ZIMBARDO, P. G., & MILLER, N. E. Facilitation of exploration by hunger in rats. *J. comp. physiol. Psychol.*, 1958, **51**, 43–46.
79. ZIMBARDO, P G., & MONTGOMERY, K. C. Effects of "free-environment" rearing upon exploratory behavior. *Psychol. Reports*, 1957, **4**, 589–594.

Received December 1, 1957.

PSYCHOLOGICAL BULLETIN
VOL. 55, NO. 5, 1958

A LEARNING THEORY APPROACH TO RESEARCH IN SCHIZOPHRENIA[1]

SARNOFF A. MEDNICK[2]
Harvard University

This paper is concerned with some aspects of schizophrenic behavior that seem amenable to interpretation in terms of learning theory. There is no pretense of completeness or uniqueness of explanation, univocal experimental support or lack of other theories to explain the same behavior. The advantage of this explanation has been the readiness with which it is tested, the success of the tests and the support which the clinical literature lends.

The discussion will take the following form. (*a*) First we will review certain experiments studying the conditioning, generalization, and learning of schizophrenics and show how these results can be understood in terms of learning theory. (*b*) We will attempt to outline some possible conditions for the onset of the thinking disorder in schizophrenia. Here we will be suggesting that generalization and high levels of anxiety may be mutually supportive and augmentative; generalization under these conditions may thus become excessive leading to difficulties in sequential thought. (*c*) The nature of the explanation of the thinking disorder will suggest the reason for the transition from the acute to the chronic phase of the illness. The thinking of abstracted, irrelevant thoughts may be rewarded by anxiety reduction by removing disturbing ideation from consciousness. This would increase the probability of the recurrence of these irrelevant thoughts and would be an admirable vehicle for continual anxiety reduction and transition to a chronic phase.

[1] This work was supported by Research Grant M-1519 from the National Institute of Mental Health.

[2] Now at the Institute of Personality Assessment and Research, University of California, Berkeley.

RESEARCH ON CONDITIONING, LEARNING, AND GENERALIZATION

The lack of anxiety of the schizophrenic, often considered an aspect of "flat affect" or emotionlessness, has received considerable attention in the clinical and experimental literature (**1, 4, 6, 30**). This reduction in reactivity has also been the focus of many theoretical discussions of the disorder. However, while "flat affect" might be a term descriptive of certain schizophrenic patients (mainly chronic, although even this has been questioned [**41**]), it has not normally been used in describing the incipient or acute patient. Thus Arietti calls the first stage of schizophrenia "a period of intense anxiety and panic." Experimental studies seem to suggest that the acute patient is very reactive showing extreme anxiety (**13, 36, 39, 61, 74, 79**). Research also suggests that such an emotionally disturbed individual will take longer to recover from an affective imbalance than will normals (**13, 14, 79**). This hyperreactivity has manifested itself in studies of heart rate, the psychogalvanic response, startle reflex, reactions to painful stimulation, etc. These studies seem to point to a low threshold of emotional arousal in at least the acute

schizophrenic. Recent work in the area of conditioning and learning has identified states of emotional arousal as contributing (as do states such as hunger and thirst) to drive strength, a construct postulated by Hull to represent the motivational force of the behavior of the organism (**19**).

All of the above may be taken to suggest that acute schizophrenics are organisms in a state of heightened drive. Hull's theory (**29**) and extensions of this theory make specific predictions of the effect of this heightened drive state on behavior. In general, the effect of heightened drive is to increase the response strength of any habit tendencies that may be aroused in a given situation. Thus, in general, the hungrier, or more anxious, or more fearful the organism, the greater his drive and the greater the speed and amplitude of his responses.

In a simple conditioning situation where the number of response alternatives is restricted (usually to one possible response) high drive would augment the response strength of the conditioned response. Thus, a group with high drive (in this case schizophrenics) should show faster conditioning than a low drive group. There have been some studies that indicate that schizophrenics can be conditioned (**5**, **25**, **37**, **51**, **59**). Other studies have compared schizophrenics with control groups for ease of conditioning. The earliest of these is the study by Pfaffman and Schlosberg (**58**). They conditioned the knee-jerk response in 25 schizophrenics and 25 normal *S*s. The schizophrenics conditioned faster. A study by Mays (**42**) and a study by Shipley (**70**) indicate that schizophrenics show faster conditioning of the psychogalvanic response than do normals. Recent studies by Taylor and Spence, and Spence and Taylor indicate that schizophrenics show faster eye blink conditioning than normals or anxiety neurotics (**73**, **76**). The results of these several studies support the above prediction that schizophrenics show faster conditioning than normals.

Stimulus generalization. When a response, having been trained to a stimulus, is also elicited by similar stimuli, stimulus generalization may be said to have occurred. As pointed out above, the effect of increased drive is to increase the response strength of aroused habit tendencies. As a consequence, high drive tends to produce heightened generalization responsiveness. It is, therefore, reasonable to expect, on the basis of these speculations, that schizophrenics would show elevated generalization responsivity.

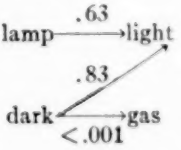
On the basis of clinical observation Schilder (**67**) first remarked on the difficulties schizophrenics have in tasks involving differentiation. Bender and Schilder studying conditioned withdrawal from shock (**5**), noted extreme over-generalization from their schizophrenic *S*s. Cameron has spoken of over-inclusion and the broadening of the generalization gradient on the part of schizophrenics (**7**, **8**, **10**). Garmezy (**21**), studying generalization along the dimension of pitch, found schizophrenics showing more generalization than normals. This effect was especially marked under conditions of stress (strong drive arousal). Mednick, using the dimension of space found that schizophrenics, especially acute patients, generalized more than normals (**44**, **45**). Dunn tested schizophrenics with social and nonsocial materials finding, with social materials, greater generalization than was shown by the control group (**17**).

The several studies cited support the prediction that schizophrenics show elevated generalization responsivity.

Complex learning. For purposes of this discussion a complex learning task will be defined as a situation in which many irrelevant and incorrect habit tendencies are aroused with which the correct response must compete. For example, in the serial learning of a word list, words other than the next correct response are potential competing responses. In this context high drive acting impartially upon correct and incorrect response tendencies will tend to push many irrelevant responses above the evocation threshold (**20**, **75**). The suprathreshold irrelevant tendencies will interfere with the correct responses, causing a relatively large number of errors and relatively slow attainment of a list criterion. Thus, in contrast to the conditioning-type situation, it is predicted that in complex response situations schizophrenics will reach a criterion more slowly than normals and with more errors. Relative to this prediction there are a considerable array of studies demonstrating poor performance by schizophrenics in complex tasks (**4**, **30**, **31**). The *E*s often report that the schizophrenics' performance is retarded by irrelevant, incorrect responses. For example, Cameron's concept of interpenetration might be understood in these terms (**9**). A paired associate verbal learning study by Hall and Crookes showed schizophrenics performing less well than normals. Hall and Crookes attribute the poor performance to extralist intrusions (**26**). As pointed out above, Hullian theory would explain this in terms of high drive indiscriminately increasing the probability of remote, irrelevant responses being evoked.

As defined above, a paired associate verbal learning task is a very complex task. However, the complexity can be minimized in several ways. Intralist similarity and intralist cross associates could be largely eliminated. Words which are already dominant associates can be chosen as the response members of the associate pairs; this would make the most probable associates correct and eliminate them as possible interfering tendencies. Such a list of paired associates would contain pairs such as "black-white" and "table-chair." While this list minimizes the possibility of competing responses, such responses are still very possible. However, in this relatively low complexity situation the present orientation would predict that the schizophrenics would learn the list *at least* as quickly as normals and with *at least* a comparable number of errors. The "at least" refers to the possibility that complexity is so reduced that the task approaches the level of a conditioning situation. If the task is of low enough complexity the schizophrenic should then perform better than the normals.

The complexity of a paired associate list may also be maximized. An example of such a list may be illustrated by the associate pairs:



While in the low complexity list we took the most probable word associate to a stimulus word and made it the correct response, in this maximum competition list we have made it an incorrect response. As noted, the probability of associating "gas"

to "dark" is less than one in a thousand (**65**), while the probability of associating "light" to "dark" is .83. Thus, "light" is a highly probable erroneous response to the stimulus "dark." In view of the low probability of the "dark-gas" pair other erroneous associates to "dark" such as "night" also have a high probability of evocation. High drive will tend to arouse relatively more erroneous associates and will give these associates relatively greater response strength. Thus schizophrenics would be predicted to make more errors and to take longer to learn a high complexity list. Mednick and DeVito (**48**) compared the learning of schizophrenics and normals on such minimum and maximum complexity lists. The schizophrenic *S*s learned the low complexity list more quickly than the normals, while the normals learned the high complexity list more quickly. The interaction of Groups ×Lists was very highly significant ($F = 258.1$ at 1, 82 *df*).

Another instance of the increasing difficulty schizophrenics have as a function of increasing task complexity is described by Hunt and Cofer. From the rich data collected by the Worcester State Hospital group (**32**) in the 1930's, Hunt and Cofer compared the difference in the performance of normals and schizophrenics in reflex time, simple reaction time, and discrimination reaction time, tasks which are graded in complexity. Hunt and Cofer comment from an analysis of this and other data " . . . we have quantitative evidence that the deficit becomes greater as the complexity of the task increases (**31**)."

Thus, three characteristics of schizophrenic behavior which reliably distinguish them from normals are: (*a*) schizophrenics more easily acquire a conditioned response; (*b*) schizophrenics show greater stimulus generalization responsiveness; (*c*) schizophrenics have great difficulty performing well in complex situations, being plagued by irrelevant, tangential associative responses competing with the adequate mode of response. However, they do *at least* as well as normals on low complexity tasks.

In the framework of the present orientation all of these behavioral characteristics would be predicted in a group with extremely high drive.

The Thinking Disorder

Bleuler considered the "disturbances of association" a primary symptom of schizophrenia from which the other disordered behaviors stem.

> It appears as if those pathways of association and inhibition, established by experience had lost their meaning and significance. Associations seem to take new pathways more easily, and thus no longer follow the old preferred ways, that is, the logical pathways indicated by past experience. . . . Especially in acute conditions of schizophrenia, one often finds so complete a fragmentation of the thinking processes that they cannot result in a complete idea or action" (**6**, p. 349–350).

Cameron and Magaret described the disorder of thinking in much the same way.

> Contradictory, competing, and more or less irrelevant responses can no longer be excluded. . . . Schizophrenic patients themselves often complain about the confusion in their talk and thinking, saying that everything seems mixed up, the words do not come as they once did, thoughts rush in and are jumbled . . . . "There are a million words," one patient said, "I can't make sentences; everything is disconnected." Another patient made several attempts to speak and then gave up; the next day she complained her thoughts had been rushing through her mind so that she could say nothing, (**11**, p. 511).

Observers differ in their theoretical interpretation of the thinking disorder. Cameron refers to desocializa-

tion (11). Kraepelin speaks of disconnection of thought (34); Goldstein (22) interprets the disorder in terms of an impairment of abstract behavior; habit deterioration is Meyer's explanation (49); Pavlov (57) understands it as a sensitization caused by overstimulation leading to a condition of cortical inhibition; Hanfmann and Kasanin refer to a loss of generalizing ability (27). Bateson uses the concept of the double bind (2); Shakow notes the importance of the inability to maintain a set (69).

While these workers may differ in their theory, they will probably acknowledge that the statements of Bleuler and Cameron and Magaret describe the behavior they seek to explain. This thought disorganization resulting from "irrelevant," "fragmented," and "competing" associations may also be understood in terms of the framework here presented. In terms of the above definition, thinking is perhaps the most complex behavior in which man can engage. There are thousands of verbal thought units that potentially compete with the single thought unit that may be demanded by a given context. (For purposes of this discussion a verbal thought unit may be informally defined as a major form class, a verb, an adjective, or noun.) These competitors may be synonyms, symbolic representatives, word associates, clang associates, etc. (12, 15, 35, 43, 54, 60, 66). Due to previous association, or similarity of meaning or sound, some thought units are probable or strong competitors. Others, more remote and irrelevant, rarely have enough response strength to reach awareness. Predictions from an extension of Hullian theory would suggest that the thinking of individuals with high drive would be disrupted by the intrusion of the remote and irrelevant thought units pushed above the threshold of awareness. The writer suggests that this action of high drive upon remote response tendencies is a major root of the disordered thinking of schizophrenics. In support of this hypothesis, a recent study by M. Mednick shows that relatively remote associations have more response strength for high anxious than for low anxious normals (43).

It may be instructive to make use of this general framework to attempt to briefly explain the possible origins of what Arietti calls the first stage of schizophrenia where the thinking disorder originally manifests itself. This description of the development of the schizophrenic break grew from attempts at explaining the conditions for the onset of illness as it was described by early acute schizophrenics. It tries to reflect their feeling of being caught in expanding, spiralling vortexes of anxiety and ideation reciprocally supporting and augmenting each other. It also tries to remember that many of these early schizophrenics will go on to become chronic. Unless the acute and chronic phases are two separate illnesses, it seems likely that a *single set of principles should explain both states.*

To begin with, the preschizophrenic is an extremely anxious individual (1, 39, 40). (The possible explanations for this may run the gamut from endocrinological to psychoanalytic to learning theories but will not be the concern of this paper.) In other words, he is an individual who strongly fears many stimulus situations. His high drive level causes him to display an abnormal amount of stimulus and associative generalization (3, 43, 46, 63, 64, 78). This means that stimulus events that

are in some way similar to the stimulus situations he has learned to fear will also tend to elicit anxiety responses. (These stimulus events include thoughts.)

Some cases learn to restrict their contact with anxiety provoking events and thoughts inasmuch as this avoidance is reinforced by anxiety reduction (**50**). Since they avoid potential anxiety-producing situations, learned anxiety never extinguishes (**72**). In most cases, this borderline adjustment continues through life and is usually termed withdrawn or schizoid.

In many cases, however, some life crisis or an untoward incident (later called the "precipitating event") interferes with this adjustment. This crisis or incident may take the form of the termination of adolescence, homosexual panic, an auto accident, the death of a loved one, or simply some morbid ruminating. In any case, the event is one which will raise the individual's anxiety level. For the highly anxious preschizophrenic, this imbalance is a serious affair which may have alarming ramifications. For one thing, the high anxious person will have a relatively large anxiety response to the precipitating incident. This will temporarily push his total drive state up to unusually high levels. One important consequence of this will be the attendant increase in the level and breadth of generalization responsivity. While he previously reacted with fear to many stimulus situations, the increase in generalization will cause a large number of new stimuli to become potential anxiety arousers. The high anxious individual finds that some of the stimuli which were once safe and comfortable for him now fill him with uneasiness. While he once felt comfortable in the presence of his superior, the incremented gradient of fear stemming from learned reactions to his father now cause him to experience discomfort.

Not only does the increase in generalization increase the number of stimuli which will arouse a fear response, it also augments the amplitude of response to the old fear producing stimuli. The low anxious individual can usually allow his anxiety level to "simmer down" by means of a good night's sleep. The high anxious person will find it hard to sleep; the pressure of the greatly increased anxiety keeps thoughts running through his mind. The thoughts will be partially cued by the ongoing anxiety state and may relate to other periods of disturbance, causing still more upset.

In summary, we find that this precipating incident has a number of uncomfortable consequences. The consequences enumerated above will tend to increase the individual's anxiety level. Unfortunately, this will, in turn, serve to again raise the level and breadth of generalization responsiveness. Again, the effect of this will be to increase the probability of his encountering a fear arousing stimulus by increasing the number of such stimuli. Also, the amount of fear that a previously adequate stimulus could arouse will be increased. Again, the individual's total anxiety level is likely to continue to rise. Subjectively, the individual will begin to feel very uncomfortable. Things which he could depend on to keep his anxiety down don't work any more. The afternoon beer does not calm anymore; the bartender now makes him feel a bit uneasy. Also, situations which made him bearably anxious now tax his ability to control himself.

Assuming continued stimulation

from the world and/or continued thinking, this reciprocal augmentation of anxiety and generalization could theoretically continue unabated until some upper physiological limit of anxiety and/or generalization is reached. Long before this point is reached, however, the behavior of the individual will become noticeably unusual. His drive level will keep thoughts racing through his mind. Many of these thoughts will be out of context or silly. His fear regarding the "craziness" and uncontrolled nature of his thoughts only serves to increase the insistence of these thoughts. The sudden lack of control he has over these thoughts will seem inexplicable. He is either "going crazy" or there is some "rational" solution for all of this. In some cases, because of past experience, a rational solution suggests itself which is compounded of elements such as X-radiation, gamma rays, radio transmitters, and the FBI. This rational solution reduces anxiety more than the thought that he is going crazy. The solution is thus reinforced, increasing the probability of its being called upon as a defense (anxiety reducer) in the presence of inexplicable anxiety or thoughts.

Meanwhile, what is happening to his thinking behavior? As the spiral of anxiety and generalization mounts, his drive level may increase to an almost insupportable degree. As this is taking place, his ability to discriminate is almost totally eclipsed by his generalization tendencies (**67**). Any unit of a thought sequence might call up some remote associate and this associate will call up still another remote associate. Clang associates based on stimulus-response generalization may be frequent. Body positions accidentally associated with any fleeting periods of anxiety reduction will tend to be continually and in some cases, continuously assumed. These will be rationalized ("If I move, evil will envelop the world.") and may be maintained for long periods of time. His speech may resemble a "word salad." He will be an acute schizophrenic with a full-blown thinking disorder.

Much of the above description of this acute process rests upon the hypothesis of the reciprocal augmentation of anxiety and generalized fear responses. There is some support for an obvious deduction from this hypothesis. The reciprocal augmentation model would predict slower extinction of a learned fear response with massed trials; the longer the time between trials the greater the opportunity for the fear response from the last trial to decay and drop the total anxiety level back to the resting state. If the trials follow each other rapidly enough, Trial 2 will come at a time when the anxiety from Trial 1 has not yet been dissipated. Thus the total drive state at Trial 2 will be the resting level plus the remaining anxiety from Trial 1. Since the total drive level will be higher on Trial 2 than it was on Trial 1, the stimulus on Trial 2 should elicit an augmented anxiety response. This Trial 2 anxiety response should take longer to decay than that of previous trials and should in the manner outlined lead to an augmented Trial 3 anxiety response, and so on. Response fatigue might dampen this spiral effect somewhat. However, studies by Howat (**28**), Mednick (**47**), and Murphy and Miller (**53**) all using noxious UCSs, have obtained little or no extinction under the conditions of massed trials and relatively rapid extinction under spaced practice.

This explication of the reciprocal augmentation process clarifies an im-

portant question. Why doesn't everybody proceed to schizophrenia after an extremely anxiety provoking event? The answer lies in three factors: the individual's original drive level, his rate of recovery from anxiety states, and the number of stimuli that elicit anxiety responses from the individual.

Highly anxious individuals will tend to have relatively strong anxiety responses. The resultant large increments in drive will take long periods to decay (**14**, **16**); during these periods drive will continue to remain in a heightened state leaving the individual prone to be swept up in the reciprocal augmentation spiral.

If an individual has an abnormally slow recovery rate from anxiety arousal, a relatively small arousal state may have a long term effect equivalent to that of a major anxiety response in a normal individual. During this extended arousal period, responsivity (including anxiety responses) will be heightened and generalization will be increased. Darrow and Heath (**14**), Wulfeck (**79**), and Cohen and Patterson (**13**) have demonstrated that schizophrenics and highly anxious persons tend to have an abnormally slow recovery rate from anxiety arousal. This relationship holds even with allowance made for the magnitude of the anxiety response.

Almost by definition, the number of stimuli which elicit anxiety will be higher for the high anxious person. Thus, he will find it relatively difficult to seek out soothing stimulation or thoughts to help mitigate his response to a crisis. Almost any place he turns will be fraught with aspects that arouse uneasiness. Many thoughts will contain elements which will generalize to anxiety provoking ideas.

Thus, the low anxious person is not likely to demonstrate a large anxiety response to fear provoking stimuli; his anxiety response will decay relatively rapidly; he will find it relatively easy to avoid further anxiety stimulation. Thus, he is not likely to get enmeshed in the anxiety-generalization spiral unless the precipating trauma is extremely fear provoking and this type of stimulation continues unabated. Behavior under conditions such as this was studied by Grinker and Spiegel (**23**, **24**) who observed schizophrenic behavior in otherwise sound individuals who were subjected to the continuous strain of combat in World War II.

It should be clear by now that high drive, slow recovery rate, and number of fear arousing stimuli are highly correlated factors. A large anxiety response will take longer to decay. People who give large anxiety responses will tend to have slower recovery rates. Under conditions of high drive, generalization will be increased making for more stimuli which will elicit anxiety. This, in turn, will produce more anxiety. However, the highly anxious individual may escape the anxiety-generalization spiral by avoiding continued stimulation or by subduing his anxiety level by means of appropriate drugs (**33**, **52**, **55**, **56**, **62**, **68**, **71**, **77**). This "treatment" will not cure the individual; nor will it prevent breakdown at some later date. However, it may prevent an acute break and lessen the danger that the condition will proceed to chronicity. The source of the anxiety and slow recovery rate must still be ascertained and if possible, removed before any lasting cure may be said to have been effected. An interesting consideration of some neurophysiological approaches to ascertaining the source of

the anxiety arousal is discussed by Malmo (**18, 38**).

The Transition from the Acute to the Chronic Phase

The spiralling process described above has its own built-in stabilizer which leads the patient to chronicity. The stabilizing process is conceived of as follows.

In the midst of this high anxiety state each time the individual responds with an anxiety-provoking thought the increment in drive will produce an increment in generalization. This may result in a highly generalized, remote, irrelevant, tangential associate. A necessary result of responding with this remote associate is the removal of the anxiety provoking thought from awareness. Thus, the remote associate will be accompanied by drive reduction which will reinforce it as a response. This will increase its probability of occurring again in the context of the antecedent anxiety provoking thought or similar thoughts. This process occurring over and over again in many contexts will provide the patient with a repertoire of anxiety-reducing though inappropriate thought responses. At first his repertoire may be limited, making for an impression of stereotypy. However, eventually (perhaps after several acute breaks) his thinking will present a varied though disorganized picture. At this point, if the patient, perceiving the disorganization, responds with the anxiety provoking thought, "I am going crazy," he can defend against it by making an immediate associative transition to an irrelevant, tangential thought or making use of a well-learned rationale such as "the radiators are broadcasting to me." This disorganized thinking will be continually self-reinforcing since it will enable him to evade anxiety provoking stimuli. Eventually it will be extremely difficult to reach his awareness for any prolonged period with material that is anxiety provoking. Since he is not dependent on the world for much of his drive reduction, he will become more and more "estranged" from it. What is especially invidious is that he will find these newly-learned techniques extremely effective and efficient in controlling anxiety. The individual need not seek special kinds of stimulation; he need not engage in elaborate rituals; he need not obtain special equipment. He simply thinks irrelevant thoughts. As the term of his illness increases and he uses these techniques more and more he will demonstrate less and less emotionality. As the patient moves into the chronic phase of the illness he may develop the "flat affect" syndrome. In some cases, the remote associations may impel to action. In others, certain actions or postures may have regularly accompanied early conditions of drive reduction and may have been incidentally learned. These postures or mannerisms will tend to be repeated and to be called on to reduce anxiety.

It may be important to note that even the chronic patient is in one sense a very anxious person. He has never had the opportunity to extinguish his prepsychotic fears. They are still elicitable; all that is required is that one break through the schizophrenic's "associative curtain."

Summary

An attempt has been made to view schizophrenia as a problem in learning theory. The research in conditioning, learning, and generalization in schizophrenia has been reviewed in these terms. It was found that this

research supported such an interpretation.

An explication of the causes of an acute schizophrenic break and the transition to chronicity was attempted.

REFERENCES

1. ARIETI, S. *Interpretation of schizophrenia.* New York: Robert Brunner, 1955.
2. BATESON, G., JACKSON, D. D., HALEY, J., & WEAKLAND, J. Toward a theory of schizophrenia. *Behav. Sci.*, 1956, **1**, 251–264.
3. BEACH, F. A. Effects of testosterone propionate on copulatory behavior of sexually inexperienced male rats. *J. comp. Psychol.*, 1942, **33**, 227–247.
4. BELLAK, L., & WILSON, ELIZABETH. On the etiology of dementia praecox; a partial review of the literature 1935–1945 and an attempt at conceptualization. *J. nerv. ment. Dis.*, 1947, **105**, 1–24.
5. BENDER, L., & SCHILDER, P. Unconditioned reactions to pain in schizophrenia. *Amer. J. Psychiat.*, 1930, **10**, 365–384.
6. BLEULER, E. *Dementia praecox or the group of schizophrenias.* New York: International University Press, 1950.
7. CAMERON, N. Reasoning regression and communication in schizophrenics. *Psychol. Monogr.*, 1938, **50**, No. 1 (Whole No. 221).
8. CAMERON, N. Schizophrenic thinking in a problem-solving situation. *J. ment. Sci.*, 1939, **85**, 1–24.
9. CAMERON, N. The functional psychoses. In Hunt, J. McV., *Personality and the behavior disorders* (2 vols.), New York: Ronald, 1944.
10. CAMERON, N. Perceptual organization and behavior pathology. In R. R. Blake, & G. V. Ramsey (Eds.), *Perception an approach to personality.* New York: Ronald, 1951.
11. CAMERON, N., & MAGARET, A. *Behavior pathology.* Boston: Houghton Mifflin, 1951.
12. COFER, C. N., & FOLEY, J. P., JR. Mediated generalization and the interpretation of verbal behavior: I. Prolegomena. *Psychol. Rev.*, 1942, **49**, 513–540.
13. COHEN, L. H., & PATTERSON, M. Effect of pain on the heart rate of normal and schizophrenic individuals. *J. gen. Psychol.*, 1937, **17**, 273–279.
14. DARROW, C. W., & HEATH, L. L. Reaction tendencies related to personality. In K. S. Lashley (Ed.). *Studies in the dynamics of behavior.* Chicago: Univer. of Chicago Press, 1932.
15. DIVEN, K. E. Certain determinants in the conditioning of anxiety reactions. *J. Psychol.*, 1937, **3**, 291–308.
16. DUFFY, E. The psychological significance of the concept of "arousal" or "activation." *Psychol. Rev.*, 1957, **64**, 265–275.
17. DUNN, W. L. Visual discrimination of schizophrenic subjects as a function of stimulus meaning. *J. Pers.*, 1954, **15**, 1950.
18. ECCLES, J. C. *The neurophysiological basis of mind.* Oxford: Clarendon, 1953.
19. FARBER, I. E. Anxiety as a drive state. In Jones, M. R. (Ed). *Nebraska Symposium on motivation.* Lincoln, Nebraska: Nebraska Univer. Press, 1954.
20. FARBER, I. E., & SPENCE, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 120–125.
21. GARMEZY, N. Stimulus differentiation by schizophrenic and normal *Ss* under conditions of reward and punishment. *J. Pers.*, 1952, **20**, 253–276.
22. GOLDSTEIN, K. The significance of special mental tests for diagnosis and prognosis in schizophrenia. *Amer. J. Psychiat.*, 1939, **96**, 575–588.
23. GRINKER, R. R., & SPIEGEL, J. P. *Men under stress.* New York: Blakiston, 1945.
24. GRINKER, R. R., & SPIEGEL, J. P. *War neurosis.* New York: Blakiston, 1945.
25. GUK, E. D. The conditioned reflex activity of schizophrenics. *Sovetak. Nevropatol.*, 1934, No. 1, 78–84. (*Psychol. Abstr.*, **9**: 717)
26. HALL, K. R. L., & CROOKES, T. G. Studies in learning impairment, I: Schizophrenic and organic patients. *J. ment. Sci.*, 1951, **97**, 725–737.
27. HANFMANN, E., & KASANIN, J. Conceptual thinking in schizophrenia. *Nerv. ment. Dis. Monogr.*, No. 67, 1942.
28. HOWAT, G. Influence of inter-trial interval during extinction on 20-minute spontaneous recovery of the conditioned eyelid response. Paper presented at Midwest Psychological Association, Chicago, 1957.
29. HULL, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.
30. HUNT, J. McV. Psychological experi-

ments with disordered persons. *Psychol. Bull.*, 1936, **33**, 1–58.

31. HUNT, J. MCV., & COFER, C. N. Psychological deficit. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald, 1944.
32. HUSTON, P. E., & SHAKOW, D. Learning capacity in schizophrenia. *Amer. J. Psychiat.*, 1949, **105**, 881–887.
33. KLINE, N., & STANLEY, A. M. Use of reserpine in a neuropsychiatric hospital. *Ann. N. Y. Acad. Sci.*, 1955, **61**, 85–91.
34. KRAEPELIN, E. *Clinical psychiatry.* New York: William Wood, 1917.
35. LACEY, J. I., SMITH, R. L., & GREEN, B. A. Use of conditioned autonomic responses in the study of anxiety. *Psychosom. Med.*, 1955, **17**, 208–217.
36. LANDIS, C., HUNT, W. A., & PAGE, J. D. Studies of the startle pattern VII. Abnormals. *J. Psychol.*, 1937, **4**, 199–206.
37. LANDKOF, B. L. Bezuslovnye i uslovnye susudistye reflexi u schizofrenikov. [Unconditioned and conditioned vascular reflexes in schizophrenia.] *Trud tsentral psi Khoevrol lust.*, 1938, **10**, 37–62.
38. MALMO, R. B. Anxiety and behavioral arousal. *Psychol. Rev.*, 1957, **64**, 276–287.
39. MALMO, R. B., & SHAGASS, C. Physiologic studies of reaction to stress in anxiety and early schizophrenia. *Psychosom. Med.*, 1949, **11**, 9–24.
40. MALMO, R. B., SHAGASS, C., BELANGER, D. J., & SMITH, A. A. Motor control in psychiatric patients under experimental stress. *J. abnorm. soc. Psychol.*, 1951, **46**, 539–547.
41. MALMO, R. B., SHAGASS, C., & SMITH, A. A. Responsiveness in chronic schizophrenia. *J. Pers.*, 1951, **4**, 359–375.
42. MAYS, L. L. Studies of catatonia, V; investigation of the perserverational tendency. *Psychiat. Quart.*, 1934, **8**, 728.
43. MEDNICK, MARTHA. Mediated generalization and the incubation effect as a function of manifest anxiety. *J. abnorm. soc. Psychol.*, 1957, **55**, 315–321.
44. MEDNICK, S. A. Distortions of the gradient of stimulus generalization related to cortical brain damage and schizophrenia. Doctoral dissertation, Northwestern Univer., 1955.
45. MEDNICK, S. A. Distortions in the gradient of stimulus generalization related to cortical brain damage and schizophrenia. *J. abnorm. soc. Psychol.*, 1955, **51**, 536–542.
46. MEDNICK, S. A. Generalization as a function of manifest anxiety and adaptation to psychological experiments. *J. cons. Psych.*, 1957, **21**, 491–494.
47. MEDNICK, S. A., COPE, ELIZABETH, & WILD, CYNTHIA. Reciprocal augmentation of anxiety and stimulus generalization. Unpublished manuscript. Biology Library, Univer. Calif., Berkeley.
48. MEDNICK, S. A., & DEVITO, R. Associative competition and verbal learning in schizophrenia. Paper read at Eastern Psychological Association, Philadelphia, April, 1958.
49. MEYER, A. Fundamental conceptions of dementia praecox. *J. nerv. ment. Dis.*, 1906, **34**, 331–336.
50. MILLER, N. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.
51. MIROLYUBOV, N. G., & UGOL, N. B. [The problem of the state of the process of excitation in schizophrenics.] *Sovetsk. Psi. Khenevrol.*, 1933, **3**, 68–82. (*Psychological Abstracts*, **9**, No. 1260.)
52. MONROE, R. R., HEATH, R. G., MICKLE, W. A., & MILLER, W. A. A comparison of cortical and subcortical brain waves in normal, barbituate, reserpine and chloropromazine sleep. *Ann. N. Y. Acad. Sci.*, 1955, **61** 56–71
53. MURPHY, J. V., & MILLER, R. E. Spaced and massed practice with a methodological consideration of avoidance conditioning. *J. exp. Psychol.*, 1956, **52**, 77–81.
54. NOBLE, C. E. Conditioned generalization of the GSR to a subvocal stimulus. *J. exp. Psychol.*, 1950, **40**, 15–25.
55. NOCE, R. H., WILLIAMS, D. B., & RAPAPORT, W. Reserpine (Serpasil) in the management of the mentally ill and mentally retarded. *J. Amer. Med. Ass.*, 1954, **156**, 821–124.
56. OWENS, J. W. M., & WALK, R. D. The effect of reserpine on avoidance learning. Paper read at Eastern Psychological Association, Atlantic City, April, 1956.
57. PAVLOV, I. P. *Conditioned reflexes.* Oxford: Oxford Univer. Press, 1927.
58. PFAFFMAN, C., & SCHLOSBERG, H. The conditioned knee jerk in psychotic and normal individuals. *J. Psychol.*, **1**, 201–206.
59. RAZRAN, G. Withdrawal responses with shock as the conditioning stimulus in

adult human *Ss. Psychol. Bull.*, 1934, **31**, 111–143.

60. RAZRAN, G. Semantic and phonetographic generalization of salivary conditioning to verbal stimuli. *J. exp. Psychol.*, 1949, **39**, 642–652.
61. RICHTER, C. P. Electrical skin resistance. Diurnal and daily variation in psychopathic and normal persons. *Arch. Neurol. Psychiat.*, 1928, **19**, 488–508.
62. RINALDI, F., & HIMWICH, H. W. A comparison of the effects of reserpine and some barbiturates on the electrical activity of cortical and subcortical structures of the brain of rabbits. *Ann. N. Y. Acad. Sci.*, 1955, **61**, 27–35.
63. ROSENBAUM, G. Stimulus generalization as a function of level of experimentally produced anxiety. *J. exp. Psychol.*, 1953, **45**, 35–43.
64. ROSENBAUM, G. Temporal gradient of response strength with two levels of motivation. *J. exp. Psychol.*, 1951, **41**, 261–267.
65. RUSSELL, W. A., & JENKINS, J. J. *The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test.* Technical report No. 11, ONR Contract N 8 onr-66216, Univer. of Minnesota, 1954.
66. RUSSELL, W. A., & STORMS, L. H. Implicit verbal chaining in paired associate learning. *J. exp. Psychol.*, 1955, **49**, 287–293.
67. SCHILDER, P. The psychology of schizophrenia. *Psychoanal. Rev.*, 1939, **26**, 380–398.
68. SCHNEIDER, J. A., & EARL, A. W. Effects of Serpasil on behavior and autonomic regulating mechanisms. *Neurol.*, 1954, **4**, 657–667.
69. SHAKOW, D. The nature of deterioration in schizophrenic conditions. *Nerv. ment. Dis. Monogr.*, No. 70, 1946.
70. SHIPLEY, W. C. Studies of Catatonia, VI: Further investigation of the perseverative tendency. *Psychiat. Quart.*, 1934, **8**, 736–744.
71. SMITH, R. P., WAGMAN, A., & RIOPELLE, A. J. Effects of reserpine on conditioned avoidance behavior in normal and brain-operated monkeys. *J. Pharm. exp. Therapeut.*, 1956, **117**, 136–141.
72. SOLOMON, R. L., & BRUSH, E. S. Experimentally derived conceptions of anxiety and aversion. In Jones, M. R. (Ed.), *Nebraska Symposium on motivation.* Lincoln, Nebraska: Nebraska Univer. Press, 1956.
73. SPENCE, K. W., & TAYLOR, JANET, A. The relation of CR strength to anxiety in normal, neurotic, and psychotic *Ss. J. exp. Psychol.*, 1953, **45**, 265–272.
74. SYZ, H. C. Psychogalvanic studies in schizophrenia. *Arch. Neur. Psychiat.*, 1926, **16**, 747–760.
75. TAYLOR, JANET A. Drive theory and manifest anxiety. *Psychol. Bull.*, 1956, **53**, 303–321.
76. TAYLOR, JANET, A., & SPENCE, K. W. Conditioning level in the behavior disorders. *J. abnorm. soc. Psychol.*, 1954, **49**, 497–502.
77. WEISKRANTZ, L., & WILSON, W. A., JR. The effects of reserpine on emotional behavior of normal and brain operated monkeys. *Ann. N. Y. Acad. Sci.*, 1955, **61**, 36–55.
78. WENAR, C. Reaction time as a function of manifest anxiety and stimulus intensity. *J. abnorm. soc. Psychol.*, 1953, **48**, 129–134.
79. WULFECK, W. H. Motor function in the mentally disordered, I and II. *Psychol. Rec.*, 1941, **4**, 271–348.

Received January 28, 1958.

PSYCHOLOGICAL BULLETIN
VOL. 55, NO. 5, 1958

COMPONENTS OF VARIANCE DUE TO ACQUIESCENCE AND CONTENT IN THE F SCALE MEASURE OF AUTHORITARIANISM[1,2]

LOREN J. CHAPMAN AND R. DARRELL BOCK[3]
University of Chicago

This paper presents a method for computing the amount of acquiescence response set variance and content variance in an opinion questionnaire of the true-false or agree-disagree format. The method is used to reanalyze the results of recently published investigations of acquiescence response set occurring in the California F Scale measure of authoritarianism (**1**). For raw data, the method uses scores on the scale as originally constituted (designated F_{pos}) as well as on scales of reversed items (F_{neg}) obtained from the same group of subjects.

Several recent studies (**2**, **3**, **4**, **5**, **7**, **9**) have looked for acquiescence response set in the F scale by means of such reversed item scales. They have used the correlation between original and reversed halves of the test to infer the presence and extent of acquiescence. The logic is that if all the systematic variance of the original F scale were content variance, i.e., free of acquiescence variance, then F_{pos} and F_{neg} should correlate positively (assuming both were scored in the content direction) at a level approaching their reliabilities. On the other hand, if all systematic variance is acquiescence, the scales should correlate negatively at a level approaching their reliabilities. The findings have been that the correlations fall between these two extremes, and from the size and direction of correlation, inferences have been made concerning the presence and relative importance of acquiescence and content variance in the scale. The present method goes beyond the correlation approach in that it permits the direct estimation of the variance components attributable to acquiescence and content and provides a significance test for each. It provides also a model in which certain properties of the original and reversed scales may be interpreted. Like the correlation approach, this method assumes the adequate reversal of the items, both in content and susceptibility to acquiescence bias.

The model for the method is contained in the following observational equations: Let x_{ik} be the score of the i-th subject on the k-th replication of the positively worded scale. Let y_{ik} be the score of the same individual on the negatively worded scale scored in the direction of content. We assume

$$F_{pos}: \quad x_{ik} = \gamma_i + \alpha_i + \epsilon_{ik} \qquad [1]$$

$$F_{neg}: \quad y_{ik} = \gamma_i - \alpha_i + \zeta_{ik} \qquad [2]$$

where

γ_i is a component of score for content,
α_i is a component of score for acquiescence, and
ϵ_{ik} and ζ_{ki} are components of error with zero means,

[1] We wish to acknowledge the suggestions of John Cotton, Northwestern University, which resulted in the improvement of an earlier draft of this paper.

[2] This study was supported by funds obtained from a grant-in-aid from the Social Science Research Committee of the University of Chicago, under a grant from the Ford Foundation in the Behavioral Sciences.

[3] Now at the University of North Carolina.

random over replications and individuals, independent of each other and the other components, and similarly distributed for all individuals.

For the sake of definiteness we will take replication to mean the readministration of structured parallel forms of the scales after a moderate interval, say two to four weeks.

It might be supposed that the variance of the positive and negative scales over individuals would be essentially equal, but this appears not to be the case for the F scale. In each of the eight studies reviewed in this paper, the variance of scores for F_{pos} is greater than for F_{neg} (see Table 1), and significantly so ($p < .01$) in six of them as determined by the test for variance differences from correlated data (**10**).

There are two possible explanations for the disparity of these variances, only one of which is plausible. Conceivably the difference could reflect a lower homogeneity of the F_{neg} scale due to inadequate reversals. However, if this were the case, the error variances computed from the Kuder-Richardson reliabilities would be higher for the F_{neg} scale. As seen in Table 1, this is not the case for those studies on which such data are available. The more reasonable explanation for the disparity of these variances is that the component of response to authoritarian content, γ_i, is positively correlated with the component for acquiescence response set, α_i. Such positive correlation would raise the variance of the F_{pos} scale because content and acquiescence would operate in the same direction to raise or lower scores. On the F_{neg} scale, they would operate in opposing directions, lowering variance. The existence of this correlation has been suggested by Leavitt, Hax, and Roche (**9**), and Chapman and Campbell (**3**, **4**), the latter supporting the interpretation by a significant correlation, found in two studies, between an independent measure of response set and a composite scale, free of response set, made up of equal numbers of F_{pos} and F_{neg} items.

If the foregoing explanation of the disparity in the F_{pos} and F_{neg} variances is accepted, it becomes possible to estimate the components of random variation over subjects attributable to content, acquiescence, and their covariation, provided the variances due to replication errors are known.

Let the means of the content and acquiescence components in the population of subjects be $\bar{\gamma}$ and $\bar{\alpha}$ respectively. Then the population mean for F_{pos} is $\bar{\gamma}+\bar{\alpha}$ and for F_{neg}, $\bar{\gamma}-\bar{\alpha}$. The variances in the population of subjects of the F_{pos} and F_{neg} scores about these means have the composition

$$\sigma_x^2 = \sigma_\gamma^2 + \sigma_\alpha^2 + 2\sigma_{\gamma\alpha} + \sigma_\epsilon^2 \qquad [3]$$

$$\sigma_y^2 = \sigma_\gamma^2 + \sigma_\alpha^2 - 2\sigma_{\gamma\alpha} + \sigma_\zeta^2 \qquad [4]$$

respectively.

Subtracting [3] and [4], we have

$$\sigma_x^2 - \sigma_y^2 = 4\sigma_{\gamma\alpha} + \sigma_\epsilon^2 - \sigma_\zeta^2.$$

If we assume further that the variances of the replication errors of the two scales are equal,

$$4\sigma_{\gamma\alpha} = \sigma_x^2 - \sigma_y^2,$$

and the unbiased estimate of $\sigma_{\gamma\alpha}$ is

$$\hat{\sigma}_{\gamma\alpha} = (s_x^2 - s_y^2)/4. \qquad [5]$$

Formula [5] estimates one half the portion of variance in the F_{pos} scale

attributable to covariation of content and acquiescence response.

Similarly, taking the covariance of [1] and [2], we have for the covariation of the F_{pos} and F_{neg} scales,

$$\sigma_{xy} = \sigma_\gamma^2 - \sigma_a^2. \qquad [6]$$

Subtracting equation [6] from [3] and solving for σ_a^2 yields

$$2\sigma_a^2 = \sigma_x^2 - \sigma_{xy} - 2\sigma_{\gamma a} - \sigma_\epsilon^2$$

and substituting for $\sigma_{\gamma a}$ from [5] gives the unbiased estimate of σ_a^2.

$$\hat{\sigma}_a^2 = \tfrac{1}{2}\left\{(s_x^2 + s_y^2)/2 - r_{xy}s_x s_y - s_e^2\right\}, \qquad [7]$$

where r_{xy} is the observed correlation between F_{pos} and F_{neg}, and s_e^2 is an estimate of the replication error variance for F_{pos}. Formula [7] estimates the portion of the F_{pos} score variance attributable to the acquiescence response set.

Finally, adding equations [3] and [6] and solving for σ_γ^2 yields

$$2\sigma_\gamma^2 = \sigma_x^2 + \sigma_{xy} - 2\sigma_{\gamma a} - \sigma_\epsilon^2,$$

and the unbiased estimate of σ_γ^2 is

$$\hat{\sigma}_\gamma^2 = \tfrac{1}{2}\left\{s_x^2 + s_y^2)/2 + r_{xy}s_x s_y - s_e^2\right\}. \qquad [8]$$

Formula [8] estimates the portion of F_{pos} score variance attributable to the content of the items.

In order to make use of these results, we must face the problem of demonstrating the equality of σ_ϵ^2 and σ_ξ^2, and obtaining s_e^2. The procedure would be to construct one or more additional F_{pos} and F_{neg} scales, parallel to the original, and use them over an appropriate period to obtain parallel form reliabilities, r_{xx} and r_{yy}, for F_{pos} and F_{neg} respectively. The error variances could then be obtained from the respective standard errors of estimate.

Unfortunately, reliabilities of this type are not available for the F scale, but plausible bounds on them may be obtained from published Kuder-Richardson Case 20 and split-half reliabilities. Replication error may be regarded as arising from three sources—(*a*) faulty item equivalence between parallel forms, (*b*) temporal variability of the subjects between test and retest, and (*c*) "within-form" error due to ambiguous items, guessing or random marking on the part of the subjects, etc. Kuder-Richardson reliabilities cover the first and third of these sources but also reflect heterogeneity of items within forms (6). Except for very homogeneous tests, (e.g., spelling tests), the contribution of within-form heterogeneity to error is probably greater than that due to temporal instability of the subjects; hence, the Kuder-Richardson reliability could be taken as a lower bound on parallel-form reliability. Furthermore, since temporal instability of the subjects is probably independent of the scale used, the equality of s_e^2 and s_ϵ^2 computed from Kuder-Richardson reliabilities could be taken as evidence that $\sigma_\epsilon^2 = \sigma_\xi^2$, as required for equation [5].

Split-half reliabilities, on the other hand, reflect primarily within-form error and only slightly item heterogeneity. They may be taken as upper bounds on the parallel form reliability when the subjects are completely stable temporally.

In Table 1, reliabilities and error variance estimates available from published studies are exhibited. It appears that a conservative estimate of the ratio of s_e^2 to the total variance, s_x^2, for F_{pos} would be .30. The most optimistic estimate of this ratio is .11, derived from the split-half reliability in the fifth study. However, some variance should be allowed for temporal instability of the subjects, so we might accept .15 as the most

TABLE 1

COMPONENTS OF VARIANCE COMPUTED FROM DIVERSE STUDIES

	Undergraduate Populations						Med. Stud.	Misc.
	Ref. (3)	Ref. (3)	Ref. (4)	Ref. (5)	Ref. (7, 8)	Bookbinder[d]	Ref. (5)	Ref. (9)[a]
N	134	144	184	152	77	179	346	133
No. of F_{pos} items	10	16	30	10	30	15	10	14
Total variances								
F_{pos}: s_x^2	32.53	7.07	19.79	.71	.45	58.50	.74	203.83
F_{neg}: s_y^2	19.51	5.31	8.98	.61	.21	31.41	.53	67.12
Reliabilities[c]								
F_{pos}	.69	.53	.71	.52	—	—	.60	.69
F_{neg}	.41	.41	.42	.43	.77	—	.42	—
Correl. of F_{pos} & F_{neg}								
r_{xy}	.17	−.01	.29	.05	−.35[b]	.17	.19	.25
Error variance								
F_{pos}: $s_e^2 = (1 - r_{xx})s_x^2$	10.08	3.32	5.74	.34	—	—	.30	63.18
F_{neg}: $s_e^2 = (1 - r_{yy})s_y^2$	11.51	3.13	5.21	.35	.05	—	.31	—
Covar. of Cont. and Acq.								
$2\sigma_{\gamma\alpha}$	20%	12%	27%	7%	27%	23%	14%	34%
Variance components								
Assuming $s_e^2 = .30s_x^2$								
Content:	32%	28%	31%	34%	10%	30%	36%	25%
Acquies.:	18%	29%	11%	29%	33%	17%	20%	11%
Assuming $s_e^2 = .15s_x^2$								
Content:	39%	36%	39%	41%	17%	37%	43%	33%
Acquies.:	26%	37%	19%	37%	42%	25%	27%	19%

[a] We are indebted to H. J. Leavitt for supplying us with the ingredients to compute some of these values.
[b] The correlation of Jackson, et al. is reversed in sign here because they did not score the F_{neg} scale in the content direction, i.e., they counted number of F_{neg} agreements rather than disagreements.
[c] Kuder-Richardson "Formula 20" reliability values are reported in References (3) and (4), correlation with 14 more F_{pos} items in (9), corrected split-half in (7, 8).
[d] Bookbinder used the reversals of Christie et al. (5). We are indebted to L. J. Bookbinder for access to his data.

favorable value for s_e^2/s_x^2. These bounds are generous enough to allow for differences in the reliabilities of the forms of the F scale used by the various investigators.

Estimates of variance attributable to content, acquiescence, and their covariance have been calculated assuming the foregoing bounds for s_e^2 and are shown in Table 1 as percentages of the total F_{pos} variances.

In interpreting the values for the variance components listed in Table 1, one should remember the importance of the assumption that the items were adequately reversed. Inspection of formulae [7] and [8] reveals that item reversals which disrupt the parallelism of the scales will distort the correlation between F_{pos} and F_{neg} towards negative values, and thus deflate the estimate of the content variance and inflate the estimate of the acquiescence variance.[4]

If it can be assumed that the replication errors for the positive and negative scale are normally distributed and their variances equal and known from an independent determination of parallel-form reliability, the statistical significance of the components of variance for content and acquiescence can be tested. If Formulae [7] and [8] are rewritten as

$$\hat{\sigma}_\gamma^2 = \tfrac{1}{2}\left\{ s_x^2 + 2r_{xy}s_x s_y + s_y^2)/2 - s_e^2 \right\} \quad [8]$$

$$\hat{\sigma}_\alpha^2 = \tfrac{1}{2}\left\{ (s_x^2 - 2r_{xy}s_x s_y + s_y^2)/2 - s_e^2 \right\} \quad [7]$$

it becomes evident that the terms involving the variances and covariances of the scales are the between-subject (Equation [8]) and within-subject mean squares (Equation [7]) in a one-way analysis of variance, corrected for the difference in the means of the positive and negative scales. The replication error of the observations is $s_e^2 = s_x^2$. In order for $\hat{\sigma}_\gamma^2$ and $\hat{\sigma}_\alpha^2$ to be unattributable to error, these mean squares must significantly exceed s_e^2. Since the mean squares are independent of s_e^2, we may define the F statistics

Content:

$$F_{N-1,n} = (s_x^2 + 2r_{xy}s_x s_y + s_y^2)/2s_e^2$$

Acquiescence:

$$F_{N-1,n} = (s_x^2 - 2r_{xy}s_x s_y + s_y^2)/2s_e^2$$

where n equals the degrees of freedom upon which the estimate of s_e^2 is based.

For the studies of Table 1, independent estimates of s_e^2 are not available. We believe, however, that $s_e^2 = .30s_x^2$ is the least favorable value for this error which could conceivably be encountered. Taking this figure as a population value, so that the F table is entered with $n = \infty$ for the lesser mean square, we find significant components of variance ($p < .01$) for content and acquiescence in all the studies.

The findings in Table 1 are varied. However, there is fairly close agreement on the content variance among the college samples with the exception of Jackson, et al. (**8**), whose findings indicate a much lower content component than do the other studies. It is best to discount the Jackson findings in drawing conclusions about the variance components because an atypically low content estimate would result from any fortuitous introduction of extraneous systematic variance, as from the use of poor item reversals. It seems safe, therefore, to conclude from Table 1 that for college populations the content variance in the F scale is around 30% to 40%, and that the remaining reliable variance divides between acquiescence and content-acquiescence covariation, in proportions which depend on the specific population sampled.

The substantial fraction of variance attributable to covariation of content and acquiescence is not entirely detrimental since it produces sizeable "valid" variance in addition

[4] Christie, at al. wrote their paper in part as a critique, on these grounds, of other reversal studies referenced in Table 1 as (**3**), (**7**), and (**9**), as well as that of Bass (**2**). Christie, et al. stated that they provided psychologically more meaningful reversals of the items than were provided in the earlier studies, and that as a result they found higher correlations between original and reversed scales, and therefore much less evidence of acquiescence response set than did the other studies. They defend the superiority of their reversals in a convincing manner. However, the present analysis of their and the earlier results fails to support their formulation of their empirical findings. Actually their correlations between the original and reversed F scales are in the same range as those reported by other investigators, when the corrections for attenuation of Christie et al. are removed. As shown in Table 1, the components of variance due to content and acquiescence computed from their data are comparable to those of the other studies.

to the extraneous acquiescence variance. The findings of this study thus lend some support to the conjecture of Leavitt, Hax, and Roche (**9**) that the "all-positive F scale is more discriminating . . . than it would have been if it included both kinds of items." Similar effects must occur in any test where response to the format is correlated with the trait measured, as, for example, in group tests of general intelligence which require reading. Acquiescence set is unique, however, in that merely by writing the items in the negative it is possible, by the method presented in this paper, to evaluate this effect quantitatively.

Whether the added acquiescence variance in the F scale is "valid" or "invalid" depends on what is being predicted. Prediction of a comparatively acquiescence free measure, such as ratings on Fascism by counselors, would be improved by use of the all positive form. But in predicting scores on another paper-and-pencil test which itself contains extraneous acquiescence variance, the all positive scale would produce too high a correlation. Conversely, a scale containing both positive and negative items would show too low a correlation. The correct degree of relationship could be estimated only by administering positive and negative forms of both tests and evaluating the component of covariance between them attributable to content, and to covariance of content and acquiescence in each.

In the study of individual differences, it may be of interest to estimate for each individual the components of his score attributable to content and acquiescence respectively. For the population of responses from individual i, the mean of half the sum of F_{pos} and F_{neg} is γ_i, and the mean of half the difference is α_i. For scores on a single occasion k, these components are estimated by

$$c_{ik} = (x_{ik} + y_{ik})/2$$

and

$$a_{ik} = (x_{ik} - y_{ik})/2,$$

respectively, where x is the score on F_{pos} and y the score on F_{neg}.

REFERENCES

1. Adorno, T. W., Frenkel-Brunswick, Else, Levinson, D. J., & Sanford, R. N. *The authoritarian personality.* New York: Harper, 1950.
2. Bass, B. M. Authoritarianism or acquiescence. *J. abnorm. soc. Psychol.*, 1955, **51**, 616–623.
3. Chapman, L. J., & Campbell, D. T. Response set in the F scale. *J. abnorm. soc. Psychol.*, 1957, **54**, 129–132.
4. Chapman, L. J., & Campbell, D. T. Acquiescence set in measures of anxiety, authoritarianism and ethnocentrism. Unpublished manuscript.
5. Christie, R., Havel, Joan, & Seidenberg, B. Is the F scale irreversible? *J. abnorm. soc. Psychol.*, 1958, **56**, 143–159.
6. Haggard, E. A. *Intraclass correlation.* New York: Dryden Press, 1958.
7. Jackson, D. N., & Messick, S. J. A note on "ethnocentrism" and acquiescent response sets. *J. abnorm. soc. Psychol.*, 1957, **54**, 132–134.
8. Jackson, D. N., Messick, S. J., & Solley, C. M. How "rigid" is the "authoritarian?" *J. abnorm. soc. Psychol.*, 1957, **54**, 137–140.
9. Leavitt, H. J., Hax, H., & Roche, J. H. "Authoritarianism" and agreement with things authoritative. *J. Psychol.*, 1955, **40**, 215–221.
10. Walker, H. M., & Lev, J. *Statistical inference.* New York: Henry Holt, 1953.

Received December 17, 1957.

PSYCHOLOGICAL BULLETIN
VOL. 55, NO. 5, 1958

MORE ON THE WILSON TEST

QUINN McNEMAR
Stanford University

Sheffield (**3**) shows that the Wilson (**4**) distribution-free test for analysis of variance hypotheses can be simplified, conceptually and computationally, by regarding the cell frequencies (e.g., the number of persons exceeding the over-all median) as "scores," then performing the usual two-way classification analysis on these scores (one per cell) with a breakdown into row, column, and $R \times C$ interaction sums of squares with *df*'s following the usual rules, *but* with the difference that one may regard the within cells variance as the variance of a binomial frequency distribution. This is a theoretical or a priori variance given (when each cell of the table of original measures contains m persons) by mpq, which under the null hypotheses of no row, no column, and no interaction effects becomes $m(.5)(.5)$ when exactly half the original scores are classifiable as being above the over-all median. This theoretical binomial variance, with its implied *df* of infinity, is used as the error term for testing row, column, and interaction effects in the analysis of the frequencies treated as scores.

Sheffield goes on to point out that an F, so obtained, when multiplied by the *df* of its numerator variance equals exactly the corresponding χ^2 of the computationally more complicated Wilson test. Since there is empirical evidence (**1**) that Wilson's test is highly erratic, it follows that its exact equivalent as given by Sheffield is also not dependable.

Actually, Sheffield's version provides a clue for seeing logically why the Wilson test (and Sheffield's modification) is unsound. Consider the Wilson-Sheffield illustrative data as set forth in Table 1, in which the entries (number out of 16 cases who are above the over-all median) are the frequency "scores" for Sheffield's two-way analysis of variance. Now it happens in this example that, even by the highly inefficient Wilson test, the row (Dial) differences are significant at the .000,001 level. Having thus not only convincingly disproved the null hypothesis but also provided evidence of a substantial Dial effect one must next consider what this does to the a priori binomial variance or error term. If all three possible null hypotheses were true, the expected frequency ("score") for each cell would be $mp = 16(.5) = 8$. About this expected value we have a variance of $mpq = 16(.5)(.5) = 4$, a value which holds for all cells under null conditions. But in this example we have definite evidence that $p = .5$ cannot hold for all the cells in rows A and C, hence the binomial error variance of 4 cannot possibly be regarded as giving the correct variance for all the cells in these two rows. The within cells variance would be of the order

TABLE 1

		Illumination			Totals
		1	2	3	
Dials	A	14	12	11	37
	B	9	7	8	24
	C	6	3	2	11
Totals		29	22	21	72

16(.25)(.75) or 3 for row A and for row C cells. This puts us in a quandary: knowing that there is a real row effect, we cannot proceed to test for either column or interaction effects because we know also that the a priori binomial error variance is inappropriately large.

The foregoing proposition is more general: as soon as we have evidence that *any* one effect is real we also have evidence that the Sheffield error term is too large for testing the other two effects—the greater any one effect, the more erroneous the error term. Furthermore, if any one effect is sizable, it follows that homogeneity of cell variances does not hold. And still another predicament: once one effect is demonstrated we know we are dealing with frequencies (Sheffield's "scores") the magnitudes of which are not independent of the variances, a fact which is analogous if not identical to violating the assumption of independence of means and variances.

Thus we have a sad state of affairs the like of which does not befuddle ordinary analysis of variance wherein tests are made without conditional considerations. Now that we know from empirical results (**1**) and from the above logical treatment that the Wilson test is definitely unsound, it may be anticlimactic to point out that Rao (**2**), Wilson's cited source for his χ^2 approach, explicitly (p. 196) mentions the difficulty to be encountered when one marginal (i.e., main) effect is found. The decomposition of χ^2, basic to the Wilson-Sheffield test, simply ceases to be possible.

REFERENCES

1. McNemar, Q. On Wilson's distribution-free test of analysis of variance hypotheses. *Psychol. Bull.*, 1957, **54**, 361–362.
2. Rao, C. R. *Advanced statistical methods in biometric research.* New York: Wiley, 1952.
3. Sheffield, F. D. Comment on a distribution-free factorial-design analysis. *Psychol. Bull.*, 1957, **54**, 426–428.
4. Wilson, K. V. A distribution-free test of analysis of variance hypotheses. *Psychol. Bull.*, 1956, **53**, 96–101.

Received December 10, 1957.